(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
15 May 2003 (15.05.2003)

PCT

(10) International Publication Number
WO 03/039443 A2

(51) International Patent Classification[7]: A61K

(21) International Application Number: PCT/EP02/12303

(22) International Filing Date: 4 November 2002 (04.11.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
01126244.1 5 November 2001 (05.11.2001) EP
02009758.0 30 April 2002 (30.04.2002) EP

(71) Applicants (*for all designated States except US*): DEUTSCHES KREBSFORSCHUNGSZENTRUM [DE/DE]; Im Neuenheimer Feld 280, 69120 Heidelberg (DE). LUDWIG-MAXIMILIANS -UNIVERSITÄT [DE/DE]; Geschwister-Scholl-Platz 1, 80539 Munich (DE).

(71) Applicants and
(72) Inventors: **HAFERLACH, Torsten** [DE/DE]; Springerstrasse 8, 81477 Munich (DE). **SCHOCH, Claudia** [DE/DE]; Springerstrasse 8, 81477 Munich (DE). **KERN, Wolfgang** [DE/DE]; Hanfelder Strasse 101, 82319 Starnberg (DE).

(72) Inventors; and
(75) Inventors/Applicants (*for US only*): **KOHLMANN, Alexander** [DE/DE]; Schwarzstrasse 14, 92318 Neumarkt (DE). **SCHNITTGER, Susanne** [DE/DE]; Saalburgstrasse 2a, 81375 Munich (DE). **DUGAS, Martin** [DE/DE]; Am Heidebruch 6, 81377 Munich (DE). **EILS, Roland** [DE/DE]; Strahlenberger Strasse 26, 69198 Schriesheim (DE). **BRORS, Benedikt** [DE/DE]; Im Linsenbühl 42, 69221 Dossenheim (DE). **MERGENTHALER, Susanne** [DE/DE]; Grubmühlerfeldstr. 19, 82131 Gauting (DE).

(74) Agent: **VOSSIUS & PARTNER**; Siebertstrasse 4, 81675 Munich (DE).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 03/039443 A2

(54) **Title:** NOVEL GENETIC MARKERS FOR LEUKEMIAS

(57) **Abstract:** The present invention is related to methods for detecting leukemia cells by determing the expression profile of a group of markers. In particular, the type or subtype of leukemia cells in a sample is determined. Further, uses of the group of markers is discloses and compositions comprising these markers.

Novel Genetic Markers for Leukemias

The present invention is related to methods for detecting leukemia cells by determing the expression profile of a group of markers. In particular, the type or subtype of leukemia cells in a sample is determined. Further, uses of the group of markers are disclosed and compositions comprising these markers.

In the present specification, a number of documents is cited. The disclosure content of these documents including manufacturers' manuals, is herewith incorporated by reference. This holds particular true for the documents such as gene accession numbers cited in Tables 43a, b, 44 and 45 providing the complete nucleotide sequence of marker genes/cDNAs. In other terms, by reciting these documents, applicant intends to incorporate the complete nucleotide/amino acid sequence of those markers where only a partial sequence has been identified in the appended Tables. It is also intended to include the (poly)peptide sequences translated from these nucleotide sequences within the disclosure content of the present specification.

Today leukemias are classified into four different groups or types: acute myeloid (AML), acute lymphatic (ALL), chronic myeloid (CML) and chronic lymphatic leukemia (CLL). Within these groups, several subcategories can be identified further using a panel of standard techniques as described below. The incidence of leukemias is increasing with age and is 5/100.000/year in AML, 1/100.000/year in ALL, 1/100.000 in CML and 6/100.000/year in CLL. Several methods for classification have to be applied at diagnosis and before treatment starts: cytomorphology and cytochemistry, multiparameter –immunophenotyping, cytogenetics including fluorescence in situ hybridization, and molecular techniques such as polymerase chain reaction (PCR). So far only a combination of these techniques allows a precise diagnosis which is necessary to apply state of the art treatment. As the exact diagnosis is mandatory for example in CML the detection of a specific cytogenetic abnormality, the translocation (9;22) or its molecular counterpart, the BCR/ABL rearrangement is required to establish the diagnosis of CML. While all patients with CML show a BCR-ABL-rearrangement and are therefore homogenous with regard to the primary genetic abnormality, in AML and

ALL at least 10-15 different subgroups have been identified on the morphological, genetical or molecular level. Also in CLL several subgroups can be clearly separated. These different subcatgories in leukemias are associated with varying clinical outcome and therefore are the basis for different treatment strategies. The importance of highly specific classification may be illustrated in detail further for the AML as a very heterogeneous group of diseases.

Data from clinical trials showed that outcome of patients with AML differs in a broad range. Several parameters influencing prognosis have been identified. These can be assigned to different categories: patients' characteristics (i.e. age, comorbidity), therapy, and biology of the AML. Therefore, a lot of effort was invested to identify biological entities and to distinguish subgroups of AML which are associated with a favorable, intermediate or unfavorable prognosis, respectively. In order to allow a comparison between different studies a classification of AML was mandatory. In 1976 the FAB classification was proposed by the French-American-British co-operative group which was based on cytomorphology and cytochemistry in order to separate AML subgroups according to the morphological appearance of blasts in the blood and bone marrow. In addition, it was recognized that genetic abnormalities occurring in the leukemic blast had a major impact on the morphological picture and even more on the prognosis. So far, the karyotype of the leukemic blasts is the most important independent prognostic factor regarding response to therapy as well as survival. For clinical purposes karyotype analysis allows to discriminate between three major prognostic groups. A favorable outcome under currently used treatment regimens with cure rates from 50 % up to 858 was observed in several studies in patients with a) t (8;21) (q22; q22) occuring in AML M2, b) inv (16) (p13q22) occurring in; AML M4eo and c) t(15;17) (q22; qll-12) occurring in AML M3/H3v. In contrast, chromosome aberrations with an unfavorable clinical course are -5/del(5q), -7/de1(7q), inv(3)/t(3:31 and complex aberrant karyotypes with cure rates of only 10%. The remainder of AML patients are assigned to a prognostically intermediate group. This latter group is very heterogeneous because it includes patients with a normal karyotype as well as those with rare chromosome aberrations with yet unknown prognostic impact.

The sub-classification of leukemias becomes increasingly important to guide therapy. Thus, the development of new, specific treatment approaches requires the identification of specific subtypes that may benefit from a distinct therapeutic protocol. It has already been shown in two entities that the development of specific

drugs can improve outcome of distinct subsets of leukemia. One important example is the development of a new therapeutic drug (STI571) for the treatment of chronic myeloid leukemia (ML): this designed molecule inhibits the CML specific chimeric tyrosine kinase BCR-ABL generated from the genetic defect observed in CML, the BCR-ABL-rearrangement due to the translocation between chromosomes 3 and 22 (t(9;22) (q34; q11)). First data show that therapy response is dramatically higher in patients treated with this new drug as compared to all other drugs that had been used so far. Another example is the subtype of acute myeloid leukemia AML M3 and its variant M3v both with karyotype t[15;17)(q22; q11-12). The introduction of a new drug (all-trans retinoic acid - ATRA) has improved the outcome in this subgroup of patient from about 50% to 85 % long-term survivors; As it is mandatory for these patients suffering from these specific leukemia subtypes to be identified as fast as possible so that the best therapy can be applied, diagnostics today must accomplish sub-classification with maximal precision. Not only for these subtypes but also for several other leukemia subtypes different treatment approaches could improve outcome. Therefore, rapid and precise identification of distinct leukemia subtypes is the future goal for diagnostics.

So far a combination of methods is necessary to obtain the most important information in leukemia diagnostics: Analysis of the morphology and cytochemistry of bone marrow blasts and peripheral blood cells is necessary to establish the diagnosis. In some cases the addition of immunophenotyping is mandatory to separate very undifferentiated AML from acute lymphoblastic leukemia and CLL. Leukemia subtypes investigated can be diagnosed by cytomorphology alone, only if an expert reviews the smears. However, a genetic analysis based on chromosome analysis, fluorescence in situ hybridization or RT-PCR and immunophenotyping is required in order to assign all cases in to the right category. The aim of these techniques besides diagnosis is mainly to determine the prognosis of the leukemia. A major disadvantage of these methods, however, is that viable cells are necessary as the cells for genetic analysis have to divide in vitro in order to obtain metaphases for the analysis. Another problem is the long time of 72 hours from receipt of the material in the laboratory to obtain the result. Furthermore, great experience in preparation of chromosomes and even more in analyzing the karyotypes is required to obtain the correct result in at least 90% of cases. These experts in their field are necessary for all other techniques

mentioned above as well. Accordingly, standard diagnosis of leukemia uses a combination of complementary methods, is expensive, time-consuming, and requires experienced experts in the field. Methods that have to be combined are cytomorphology or histomorphology, multiparameter-immunophenotyping, cytogenetics, fluorescence in situ hybridization, and molecular genetics such as polymerase chain reaction based assays.

Using these techniques in combination, hematological malignancies in a first approach are separated into chronic myeloid leukemia (CML), chronic lymphoid (CLL), acute lymphoblastic (ALL), and acute myeloid leukemia (AML). Within the latter three disease entities several prognostically relevant subtypes have been established. As a second approach this further subclassification is based mainly on genetic abnormalities of the leukemic blasts and clearly is associated with different prognoses. Therefore, this subclassification is increasingly important to guide therapy. Furthermore, the development of new, specific treatment approaches requires precise identification of leukemia subtypes.

In a first study Golub et al. (Science 1999) showed that gene expression profiles can be used for class prediction and discriminated AML from ALL samples. However, for his analysis of acute leukemias the selection of the two different subgroups was performed using exclusively morphologic-phenotypical criteria. This was only descriptive and does not provide deeper insights into the pathogenesis or the underlying biology of the leukemia. The approach reproduces only very basic knowledge of cytomorphology and intends to differentiate classes. The data is not sufficient to predict prognostically relevant cytogenetic aberrations.

Thus, the technical problem underlying the present invention was to provide means for leukemia diagnostics which overcome the disadvantages of the prior art diagnostic methods.

The solution to said technical problem is achieved by providing the embodiments characterized in the claims. Accordingly, the present invention relates to a method of determining whether a patient sample contains leukemia cells or other cells comprising the steps of a) determining the expression profile of a group of markers in a patient sample and b) concluding from the expression profile whether the

patient sample contains leukemia cells or other cells characterized in that the group of markers consists of markers selected independently from the markers listed in one or more of the tables 3 to 6, tables 15 to 20, tables 29, 30, 41, or 42 and whereby the number of markers in the group is between one and the total number of markers listed in the tables 3 to 6, tables 15 to 20, and tables 29, 30, 41, or 42. In a particular embodiment therof, the present invention pertains to a method wherein leukemia type and subtype are simultaneously determined whereby a microarray for the detection of the expression level of a marker or a group of markers is used.

It is important to note that in accordance with the invention in all pertaining embodiments any possible combination of markers, said markers being disclosed in the respective table or tables is encompassed within the scope of the invention.

As used herein, the term "expression" refers to the process by which mRNA or a polypeptide is produced based on the nucleic acid sequence of a gene. The process includes both transcription and translation, i.e. „expression" shall also include the formation of mRNA upon transcription.

In accordance with the present invention, the term „determining the expression profile" preferably refers to the determination of the level of expression, namely of said group of markers.

As used herein, the <u>term</u> „marker" refers to a DNA, in particular cDNA, or RNA or a fragment thereof or a protein or a fragment thereof which are in the case of RNA (or cDNA) formed upon transcription of a nucleotide sequence which is capable of expression. The nucleic acid molecule fragments refer to fragments preferably of at least 8 such as ten, twelfe, fifteen or eighteen nucleotides in length representing a consecutive stretch of nucleotides of a gene, cDNA or mRNA such as of 20 or 25 nucleotides that are, for example, further specified in the appended Tables or a complementary sequence thereto. In other terms, markers include any fragment (or complementary sequence thereto) of the sequences depicted in the appended tables as long as these fragments unambiguously identify the marker. Typical fragment lengths are provided above. The determination of the expression profile of markers may be effected at the transcriptional or translational level. In other terms, the method of the invention envisages the determination at the level of mRNA or at the protein level. Protein fragments such as peptides advantageously

comprise at least 6 consecutive amino acids representative of the corresponding full length protein. 6 amino acids are generally recognized as the lowest peptidic stretch giving rise to a linear epitope recognized by an antibody, fragment or derivative thereof. Alternatively, the proteins or fragments thereof may be analysed using nucleic acid molecules specifically binding to three-dimensional structures (aptamers). In principle, the investigator may determine, in accordance with the method of the invention, whether a gene is expressed at all in a leukemic or other cell. Alternatively, an investigator may determine the difference in the expression level, for example, between a leukemic and a non-leukemic cell or between two or more different types or subtypes of leukemia. If the sample comprises only other, i.e. non-leukemia cells, then the patient's suffering from a leukaemia may safely be denied. Insofar, the above main embodiment is to be understood that if the presence of other cells is determined then this determination includes an assessment to the effect that only other cells but no leukemic cells are comprised in the sample. On the other hand, the determination of leukemic cells may include the further characterization of such cells including the differentiation status of the cells as well as the distinction from other types of cancer cells or other subtypes of leukaemia cells. Particular embodiments in this regard are further outlined herein below.

In accordance with the above, the present invention also contemplates methods where simply the assessment of leukaemia cells but not necessarily of other cells is effected. This holds true for all embodiments where the determination of other cells is mentioned. It is to be understood that with the exception of the possible determination of other cells, the steps of the various methods of the invention remain unchanged. Thus, the invention also relates to a method of determining whether a patient sample contains leukemia cells comprising the steps of a) determining the the expression profile of a group of markers in a patient sample and b) concluding from expression profile whether the patient sample contains leukemia cells characterized in that the group of markers consists of markers selected independently from the markers listed in one or more of the tables 3 to 6, tables 15 to 20, tables 29, 30, 41, or 42 and whereby the number of markers in the group is between one and the total number of markers listed in the tables 3 to 6,

tables 15 to 20, and tables 29, 30, 41, or 42. Thus, the invention further relates to a method of determining whether a patient sample contains leukemia cells and at the same time or subsequently determining the type and subtype of leukemia cells, if leukemia cells are present, comprising the steps of a) determining the expression profile of a group of markers in a patient sample and b) concluding from the expression profile whether the patient sample contains leukemia cells and at the same time or subsequently determining the type and subtype of leukemia cells, if leukemia cells are present, characterized in that the group of markers consists of markers selected independently from the markers listed in one or more of the tables 16 to 20 or table 29 or 30 and whereby the number of markers in the group is between one and the total number of markers listed in the tables 16 to 20 or table 29 or 30, to name two important embodiments of the invention.

Determination of the expression profile/levels may be effected by a variety of methods, depending on the nature of the marker. Thus, if the marker is mRNA, cDNA may be prepared into which a detectable label, such as a fluorescent, chemiluminescent, bioluminescent, radioactive (such as ^{3}H or ^{32}P) label is incorporated. Said detectably labelled cDNA, in single-stranded form, may then be hybridised, preferably under stringent or highly stringent conditions to a panel of single-stranded oligonucleotides representing different genes and affixed to a solid support such as a chip. Upon applying appropriate washing steps, those cDNAs will be detected or quantitatively detected that have a counterpart in the oligonucleotide panel. Various advantageous embodiments of this general method are feasible. For example, the mRNA or the cDNA may be amplified wherein it is, for quantitative assessments, preferable that the number of amplified copies corresponds relative to further amplified mRNAs or cDNAs to the number of mRNAs originally present in the cell. Also, the cDNAs may be transcribed into cRNAs wherein only in the transcription step a label is incorporated into the nucleic acid and wherein the cRNA is employed for hybridisation. Alternatively, the lable may be attached subsequent to the transcription step. Similarly, proteins from a cell or tissue under investigation may be contacted with a panel of aptamers or of antibodies or fragments or derivatives thereof. The antibodies etc. may be affixed to a solid support such as a chip. Binding of proteins indicative of a

leukemia or a subtype of leukaemia may be verified by binding to a detectably labelled secondary antibody or aptamer. For the labelling of antibodies, it is referred to Harlow and Lane, "Antibodies, a laboratory manual", CSH Press, 1988, Cold Spring Harbor. As regards further test assays and formats, it is referred to further embodiments of the invention as specified herein below as well as to the appended examples. In addition, a number of applicable assay formats are available in the art that can applied to the method of the invention without further ado. Specifically, a minimum set of proteins necessary for diagnosis of all leukemia types may be selected for creation of a protein array system to make diagnosis on a protein lysate of a diagnostic bone marrow sample directly. Protein Array Systems for the detection of specific protein expression profiles already are available (for example: Bio-Plex, BIORAD, München, Germany). For this application preferably antibodies against the proteins have to be produced and immobilized on a platform e.g. glasslides or microtiterplates. The immobilized antibodies can be labeled with a reactant specific for the certain target proteins as discussed above. The reactants can include enzyme substrates, DNA, receptors, antigens or antibodies to create for example a capture sandwich immunoassay.

The level of the expression of the „marker" is indicative of a leukemic condition, of a cell or an organism. The level of expression of a marker or group of markers is measured and is compared with the level of expression of the same marker or the same group of markers from other cells or samples. The comparison may be effected in an actual experiment or in silico. When the expression level also referred to as expression pattern or expression signature (expression profile) is measurably different, there is according to the invention a meaningful difference in the level of expression. Preferably the difference at least is 5 %, 10% or 20%, more preferred at least 50% or may even be as high as 75% or 100%. More preferred the difference in the level of expression is at least 200%, i.e. two fold, at least 500%, i.e. five fold, or at least 1000%, i.e. 10 fold.

The present invention allows to diagnose a wide variety and at least 14 different clinically relevant leukemia subtypes. Therefore, the invention of a combination of marker genes and their specific expression level it is possible to substitute all other mandatory diagnostic approaches including the approach of Golub and colleagues (cytomorphology or histomorphology, multiparameter-immunophenotyping,

cytogenetics, fluorescence in situ hybridization, and molecular genetics) in one single step with a specifity and sensitivity that had never been achieved in all other techniques used so far.

In more detail, based on biomathematical analysis of gene expression profiles a new method could be provided which forms the basis for designing and developing a novel diagnostic approach preferably based on microarray technology. Further, subsets of markers, preferably genes could be introduced which allow the determination of leukemia type and subtype. The method according to the invention abolishes today's standard procedures in diagnosis of leukemia. These standard diagnostic procedures require more and more centralized core facilities with both personal experts in the fields of cytomorphology, cytogenetics and molecular genetics and expensive lab equipment, which causes increasing costs for adequate diagnosis. The present invention provides novel cost-effective methods and diagnostic tools, which are less time consuming, easy to operate but nevertheless as accurate and safe as all standard methods combined today. The genes or sets of genes allows to assign clinical samples either as healthy or malignant simply based on their gene expression profiles. The genes, representative fragments thereof or transcription or translation products thereof form the basis for the methods of the invention or diagnostic tools, corresponding thereto. Furthermore, these genes etc. allow to predict the diagnoses based on the genetic abnormality of the expression pattern and to discriminate between different prognostic relevant entities. When comparing two groups of microarray experiments, Golub's method (Science 286 (1999), 531-537) sorts the genes with respect to the signal-to-noise ratio of gene x: $S_x = (\mu_1-\mu_2)/(\sigma_1+\sigma_2)$, where μ_k and σ_k denote the mean expression and standard deviation of gene x in group k.

According to a specified number of "informative" genes the 20 best discriminating genes are selected. For each informative gene a decision limit is calculated as $b_x = (\mu_1+\mu_2)/2$. To classify a new sample of an independent test set, the gene expression levels of informative genes are taken and for each gene x and sample y a so-called vote is calculated as $V_x = S_x\ (g_x^y - b_x)$, where g_x^y denotes expression level of gene x in sample y. The votes of all informative genes are summed up ("weighted voting") and depending upon the sign of this sum the new sample is

classified as group 1 or group 2. The *confidence* in the prediction is calculated as $|\Sigma V_x / \Sigma |V_x||$.

To assess the significance of each gene, a permutation test is performed, which determines signal-to-noise ratios when class labels are permuted randomly. To assess the robustness of the classifier, a leave-one-out crossvalidation is performed. *Accuracy* is the rate of correctly classified test samples.

The decision limit proposed by Golub does not provide optimal classification accuracy in all situations. When the standard deviation of expression levels within the two groups are very different, the decision limit is biased towards the group with the higher standard deviation.

A decision limit for a particular gene can be considered optimal, if it achieves maximum classification accuracy for a given dataset. By determining systematically classification accuracies for a set of possible decision limits, an optimal decision limit can be calculated. The underlying statistics as described in Example 3 select an optimal decision limit from the following set of decision limits L_x:

$$L_x = \{ (g_x^y + g_x^{y-1})/2 \mid 1 < y <= n \}$$

where g_x^y denotes expression level of gene x in sample y, n denotes the total number of samples in the training set.

Golubs method selects an arbitrary number of "informative" genes to discriminate between two classes of samples according to their signal-to-noise ratio, typically in the range of 10 to 50 genes.

Choosing too many genes like in Golub's method carries the risk of overfitting, which causes poor generalization features of the model.

Therefore the present invention applies an heuristic approach to select a minimal set of discriminative genes, which provides maximum classification accuracy in

leave-one-out-crossvalidation. I.e. for a given set of genes weighted voting as described by Golub is applied and the classification accuracy is calculated by crossvalidation used in accordance with the present invention and representing a further embodiment in accordance with this invention.

The method for achieving this used in accordance with the present invention and representing a further embodiment in accordance with this invention consists of the following steps:

(a) calculating of the top 20 discriminating genes according to the signal-to-noise ratio (top 20 SNR's);

(b) calculating classification accuracy and confidence based on optimal decision limits for each of the top 20 genes;

(c) selecting the gene which provides best classification accuracy and confidence out of step 2; and

(d) testing for each of the remaining 19 genes, whether adding this gene to the model improves accuracy and confidence.

If the gene improves accuracy and confidence, it is added to the weighted voting model, otherwise it is discarded.

Preferably, the decision limit is set according to the formula recited above.

In a pilot study consisting of 103 Affymetrix Genechip microarrays with 12625 genes each as shown in the appended examples we compared the results achieved with Golub's method and with our extended method.

Table A presents an analysis of 18 samples class A versus 85 samples class non-A. Based on 20 informative genes Golub's method results in a crossvalidation accuracy of 0,87 (confidence 0,77); achieves with three genes out of the top 20 set a crossvalidation accuracy of 0,96 (confidence 0,88).

The same analysis was performed for one versus all (OVA) and all pairs (AP) comparisons in this dataset consisting of 5 different classes. Figure 13 b presents accuracy and confidence obtained by both methods: the method of the invention outperforms Golub's method clearly both in terms of accuracy and confidence of classifications.

The development of a leukemia diagnostic tool, preferably microarray based, allows for all patients which are preferably humans and specimens a reproducible, highly specific and rapid method to obtain important information for treatment strategies in leukemia. This technique can be established in every laboratory using basic methods of molecular biology, and preferably makes use of hybridization and amplification such as PCR or LCR based techniques and does not require hematologists or cytogeneticists with several years of experience in leukemia diagnostics. Material for the analysis can be sent over large distances as it is not necessary that cells arrive viable in the laboratory. Therefore, a centralization of leukemia diagnostics with very high quality is possible.

Moreover, the accumulation of an immense knowledge about gene expression profiles in leukemia types and subtypes, which are not characterized by specific genetic abnormalities, leads to a more precise classification compared to all other methods used so far. In addition, the data compiled in accordance with the invention are helpful for the understanding of the pathogenesis of leukemia and will allow to identify genes which are specifically dysregulated. They may be considered as potential targets for therapeutic interventions specifically designed for the different leukemia subtypes.

Preferably the method according to the invention is characterized in that the group of markers consists of between two, such as three, four, five, six, seven, eight, nine or ten and the total number of markers listed in one or more of the tables 3 to 6, tables 15 to 20, and tables 29, 30, 41, or 42. Most preferred, the group consists of all markers listed in one or more tables, whereby the tables are selected from the the tables 3 to 6, tables 15 to 20, and tables 29, 30, 41, or 42. The invention also contemplates that all markers in all tables are analysed. This holds true for the presently discussed as well as for embodiments discussed further below.

Another embodiment of the invention relates to a method of determining whether a patient sample contains leukemia cells or other cells and at the same time or subsequently determining the type and subtype of leukemia cells, if leukemia cells are present, comprising the steps of determining the expression profile, preferably the level of expression of a group of markers in a patient sample and concluding from the (altered) expression profile i.e. the difference in the level of expression, whether the patient sample contains leukemia cells or other cells and at the same time determining the type and subtype of leukemia cells, if leukemia cells are present, characterized in that the group of markers consists of markers selected independently from the markers listed in one or more of the tables 16 to 20 or table 29 or 30 and whereby the number of markers in the group is between one, preferably two such as three, four, five, six, seven, eight, nine or ten and the total number of markers listed in one or more of the tables 16 to 20 or table 29 or 30. It is preferred that the group of markers consists of all markers listed in one or more tables, whereby the tables are selected from the tables 16 to 20 or table 29 or 30. In a preferred embodiment it is differentiated between four types of leukemia cells and the other cells in the patient sample. The other cells are preferably normal cells.

The "other cells" may be, for example, cells affected by a disease which is not a leukaemia. It is preferred, in accordance with the present invention that said other cells are normal cells, i.e. cells not affected by any disease.

This embodiment of the present invention allows for the differentiation between four different types of leukemias, i.e. AML, CLL, CML and ALL. As has been surprisingly demonstrated in accordance with the present invention, the qualitative and/or quantitative determination of an expression profile of a number of genes allows the unambiguous classing with any of the above and currently established types of leukemias. In principle and more preferred, the relation of the gene expression profile to the leukaemia type may take place at the same time at which the determination of the leukaemia cells in the sample takes place. Alternatively, the classification may be effected at a later time point. It was surprising that the distinction between the large number of leukemia types and subtypes, including the cytogenetically and immunophenotypically defined, as well as types

characterized by complex chromosomal aberations, could be accomplished preferably by the use of a microarray for the detection of the expression level of a marker or a group of markers with such ease and accuracy. In particular, certain preferred subsets of genes are provided which can either be used to determine the leukemia type and subtype, or only determine the subtypes of a certain leukemia type or differentiates certain types or subtypes, respectively, from one another.

In another embodiment a method is disclosed which allows differentiating between two types of leukemia cells or one type of leukemia cells and normal cells or non-leukemia cells in a patient sample comprising the steps of determining the expression profile preferably the level of expression, of a group of markers in the patient sample and concluding from the (altered) expression profile, i.e. the difference in the level of expression, which type of leukemia cells the patient sample contains or whether it contains (only) normal cells or non-leukemia cells characterized in that the group of markers consists of markers selected independently from the markers listed in one or more of the tables 3 to 6 or tables 7 to 12 and whereby the number of markers in the group is between one, preferably two such as three, four, five, six, seven, eight, nine or ten and the total number of markers listed in one or more of the tables 3 to 6 or tables 7 to 12. In a preferred embodiment the group of markers consists of all markers listed in one or more of the tables 3 to 6 or tables 7 to 12.

In another embodiment of the invention a method is disclosed allowing the differentiation between the subtypes of AML cells or between the subtypes of AML cells and normal cells in a patient sample comprising the steps of determining the expression profile, preferably the level of expression of a group of markers in the patient sample and concluding from the the (altered) expression profile, i.e. the difference in the level of expression, which subtypes of AML cells the patient sample contains or whether it contains normal cells characterized In that the group of markers consists of markers selected independently from the markers listed in one or more of the tables 1, 2, 13, 14, 17, 25, 27, 35 and 36 and whereby the number of markers in the group is between one, preferably two such as three, four, five, six, seven, eight, nine or ten and the total number of markers listed in one or more of the tables 1, 2, 13, 14, 17, 25, 27, 35 and 36. In a preferred

embodiment the group of markers consists of all markers listed in one or more of the tables 1, 2, 13, 14, 17, 25, 27, 35 and 36. It is preferred that three, four or more subtypes of AML cells are determined.

In another embodiment of the invention a method is disclosed allowing the differentiation between and thus the determination of the subtypes of ALL cells in a patient sample comprising the steps of (a) determining the level of expression of a group of markers in the patient sample and (b) concluding from the differences in the level of expression which subtypes of ALL cells the patient sample contains whereby the group of markers consists of markers selected independently from the markers listed in one or more of the tables 18, 32 or 33 and whereby the number of markers in the group is between one, preferably two such as three, four, five, six, seven, eight, nine or ten and the total number of markers listed in one or more of the tables 18, 32 or 33. It is preferred that the group of markers consists of all markers listed in one or more of the tables 18, 32 or 33.

In another embodiment of the invention a method is disclosed allowing the differentiation between and thus the determination of the subtypes of CLL cells in a patient sample comprising the steps of determining the level of expression of a group of markers in the patient sample and concluding from the differences in the level of expression which subtypes of CLL cells the patient sample contains whereby the group of markers consists of markers selected independently from the markers listed in one or more of the tables 38 or 39 and whereby the number of markers in the group is between one, preferably two such as three, four, five, six, seven, eight, nine or ten and the total number of markers listed in one or more of the tables 38 or 39. It is preferred that the group of markers consists of all markers listed in one or more of the tables 38 or 39.

In another embodiment of the invention, a method is disclosed of assessing the efficacy of a test compound for inhibiting leukemia, the method comprising comparing the expression profile of a group of markers in a first sample obtained from the patient and maintained in the presence of the test compound and the expression profile of a group of markers in a second sample obtained from the

patient and maintained in the absence of the test compound, wherein a significantly altered expression profile of the group of markers in the first sample, relative to the second sample, is an indication that the test compound is efficacious for inhibiting leukemia in the patient characterized in that the group of markers consists of markers selected independently from the markers listed in one or more of the tables 1 to 20, tables 25 or 27 or tables 29, 30, 32, 33, 35, 36, 38, 39, 41, 42 and whereby the number of markers in the group is between one, preferably two such as 3, 4, 5, 6, 7, 8, 9 or 10 and the total number of markers listed in the tables 1 to 20, tables 25 or 27 or tables 29, 30, 32, 33, 35, 36, 38, 39, 41, 42.

In accordance with this embodiment of the present invention, it is again preferred that in the comparison of expression profiles expression levels and differences in expression levels are determined and compared. It is further preferred that the alteration determined in accordance with the method of the invention in the expression profile or expression level must be in the direction of the expression profile of normal cells or at least diseased but non-leukemic cells. More preferably the alteration should be in the direction of normal blood cells, more preferably cells of the certain type. Accordingly, it is also preferred that the comparison includes an internal standard of expression levels of analysed markers wherein the internal standard represents the expression profile of non-leukemic and preferably normal cells. The comparison may be effected by relying on actual experimental data or on in silico obtained reference data.

In another embodiment of the invention a method is disclosed of assessing the efficacy of a therapy for inhibiting leukemia in a patient, the method comprising comparing the expression profile, preferably the level of expression of a group of markers in the first sample obtained from the patient prior to providing at least a portion of the therapy to the patient and the expression profile, preferably the level of expression of a group of markers in a second sample obtained from the patient following provision of the portion of the therapy, wherein a significantly (altered) expression profile, i.e. a significantly (altered) difference in the level of expression of the group of markers in the second sample, relative to the first sample, is an

indication that the therapy is efficacious for inhibiting leukemia in the patient characterized in that the group of markers consists of markers selected independently from the markers listed in one or more of the tables 1 to 20, tables 25 or 27 or tables 29, 30, 32, 33, 35, 36, 38, 39, 41, 42 and whereby the number of markers in the group is between one, preferably two such as 3, 4, 5, 6, 7, 8, 9 or 10 and the total number of markers listed in the tables 1 to 20, tables 25 or 27 or tables 29, 30, 32, 33, 35, 36, 38, 39, 41, or 42.

As with the previous embodiment, the alteration determined in accordance with the method of the invention in the expression profile or expression level must be in the direction of the expression profile or normal cells or at least diseased but non-leukemic cells. Accordingly, it is also preferred in accordance with this embodiment that the comparison includes an internal standarad of expression levels of analysed markers wherein the internal standarad represents the expression profile of non-leukemic and preferably normal cells. The comparison may – again – be effected by relying on actual experimental data or on in silico obtained reference data.

Within the therapy of the patient, compounds may be administered that have at least passed phase II and preferably are whithin phase III of clinical trials. Advantageously, in one embodiment, a therapeutical composition or medicinal product is administered that comprises one pharmaceutically active compound. In alternative embodiments, pharmaceutical compositions or medicinal products are administered that comprise more than one pharmaceutically active compound. If the composition or product comprises more than at least one pharmaceutically active compound then one of the compounds may aim at the direct reduction of tumor load wherein at least one further compound may fulfil an accessory function such as the general stimulation of the immune system. Compounds of the latter class are also well known in the art and comprise plant derived products as well as immunostimulatory molecules selected from the group of interleukins, interferons and others.

Additionally, the invention contemplates a method of refining a compound identified by the method as described herein above, said method comprising optionally the steps of said methods and:

(1) identification of the binding sites of the compound and the target molecule by site-directed mutagenesis or chimeric protein studies;

(2) molecular modeling of both the binding site of the compound and the binding site of the target molecule; and

(3) modification of the compound to improve its binding specificity for the target.

The target may in accordance with the above be DNA, mRNA or protein. All techniques employed in the various steps of the method of the invention are conventional or can be derived by the person skilled in the art from conventional techniques without further ado. Thus, biological assays based on the herein identified nature of the proteins/(poly)peptides may be employed to assess the specificity or potency of the drugs wherein the increase of one or more activities of the proteins/(poly)peptides may be used to monitor said specificity or potency. Steps (1) and (2) can be carried out according to conventional protocols. A protocol for site directed mutagenesis is described in Ling MM, Robinson BH. (1997) Anal. Biochem. 254: 157-178. The use of homology modeling in conjunction with site-directed mutagenesis for analysis of structure-function relationships is reviewed in Szklarz and Halpert (1997) Life Sci. 61:2507-2520. Chimeric proteins are generated by ligation of the corresponding DNA fragments via a unique restriction site using the conventional cloning techniques described in Sambrook (1989), loc. cit.. A fusion of two DNA fragments that results in a chimeric DNA fragment encoding a chimeric protein can also be generated using the gateway-system (Life technologies), a system that is based on DNA fusion by recombination. A prominent example of molecular modeling is the structure-based design of compounds binding to HIV reverse transcriptase that is reviewed in Mao, Sudbeck, Venkatachalam and Uckun (2000). Biochem. Pharmacol. 60: 1251-1265.

For example, identification of the binding site of said drug by site-directed mutagenesis and chimerical protein studies can be achieved by modifications in the (poly)peptide primary sequence that affect the drug affinity; this usually allows to precisely map the binding pocket for the drug.

As regards step (2), the following protocols may be envisaged: Once the effector site for drugs has been mapped, the precise residues interacting with different parts of the drug can be identified by combination of the information obtained from mutagenesis studies (step (1)) and computer simulations of the structure of the binding site provided that the precise three-dimensional structure of the drug is known (if not, it can be predicted by computational simulation). If said drug is itself a peptide, it can be also mutated to determine which residues interact with other residues in the (poly)peptide of interest.

Finally, in step (3) the drug can be modified to improve its binding affinity or ist potency and specificity. If, for instance, there are electrostatic interactions between a particular residue of the (poly)peptide of interest and some region of the drug molecule, the overall charge in that region can be modified to increase that particular interaction.

Identification of binding sites may be assisted by computer programs. Thus, appropriate computer programs can be used for the identification of interactive sites of a putative inhibitor and the (poly)peptide by computer assisted searches for complementary structural motifs (Fassina, Immunomethods 5 (1994), 114-120). Further appropriate computer systems for the computer aided design of protein and peptides are described in the prior art, for example, in Berry, Biochem. Soc. Trans. 22 (1994), 1033-1036; Wodak, Ann. N. Y. Acad. Sci. 501 (1987), 1-13; Pabo, Biochemistry 25 (1986), 5987-5991. Modifications of the drug can be produced, for example, by peptidomimetics and other inhibitors can also be identified by the synthesis of peptidomimetic combinatorial libraries through successive chemical modification and testing the resulting compounds. Methods for the generation and use of peptidomimetic combinatorial libraries are described in the prior art, for example in Ostresh, Methods in Enzymology 267 (1996), 220-234 and Dorner, Bioorg. Med. Chem. 4 (1996), 709-715. Furthermore, the three-dimensional and/or crystallographic structure of activators of the expression of the (poly)peptide of the invention can be used for the design of peptidomimetic

activators, e.g., in combination with the (poly)peptide of the invention (Rose, Biochemistry 35 (1996), 12933-12944; Rutenber, Bioorg. Med. Chem. 4 (1996), 1545-1558).

In accordance with the above, in a preferred embodiment of the method of the invention said at least one compound is further refined by peptidomimetics.

The invention furthermore relates to a method of modifying a compound identified or refined by the method as described herein above as a lead compound to achieve (i) modified site of action, spectrum of activity, organ specificity, and/or (ii) improved potency, and/or (iii) decreased toxicity (improved therapeutic index), and/or (iv) decreased side effects, and/or (v) modified onset of therapeutic action, duration of effect, and/or (vi) modified pharmakinetic parameters (resorption, distribution, metabolism and excretion), and/or (vii) modified physico-chemical parameters (solubility, hygroscopicity, color, taste, odor, stability, state), and/or (viii) improved general specificity, organ/tissue specificity, and/or (ix) optimized application form and route by (i) esterification of carboxyl groups, or (ii) esterification of hydroxyl groups with carbon acids, or (iii) esterification of hydroxyl groups to, e.g. phosphates, pyrophosphates or sulfates or hemi succinates, or (iv) formation of pharmaceutically acceptable salts, or (v) formation of pharmaceutically acceptable complexes, or (vi) synthesis of pharmacologically active polymers, or (vii) introduction of hydrophylic moieties, or (viii) introduction/exchange of substituents on aromates or side chains, change of substituent pattern, or (ix) modification by introduction of isosteric or bioisosteric moieties, or

(x) synthesis of homologous compounds, or (xi) introduction of branched side chains, or (xii) conversion of alkyl substituents to cyclic analogues, or (xiii) derivatisation of hydroxyl group to ketales, acetales, or (xiv) N-acetylation to amides, phenylcarbamates, or (xv) synthesis of Mannich bases, imines, or (xvi) transformation of ketones or aldehydes to Schiff's bases, oximes, acetales, ketales, enolesters, oxazolidines, thiozolidinesor combinations thereof; said method optionally further comprising the steps of the above described methods.

The various steps recited above are generally known in the art. They include or rely on quantitative structure-action relationship (QSAR) analyses (Kubinyi, "Hausch-Analysis and Related Approaches", VCH Verlag, Weinheim, 1992), combinatorial biochemistry, classical chemistry and others (see, for example, Holzgrabe and Bechtold, Deutsche Apotheker Zeitung 140(8), 813-823, 2000).

The invention moreover relates to a method of producing a pharmaceutical composition comprising optionally the steps of the aforementioned methods and further the step of formulating the at least one compound identified, refined or modified by the method of any of the preceding embodiments with a pharmaceutically active carrier or diluent.

The pharmaceutical composition produced in accordance with the present invention may further comprise a pharmaceutically acceptable carrier and/or diluent and/or excipient. Examples of suitable pharmaceutical carriers are well known in the art and include phosphate buffered saline solutions, water, emulsions, such as oil/water emulsions, various types of wetting agents, sterile solutions etc. Compositions comprising such carriers can be formulated by well known conventional methods. These pharmaceutical compositions can be administered to the subject at a suitable dose. Administration of the suitable compositions may be effected by different ways, e.g., by intravenous, intraperitoneal, subcutaneous, intramuscular, topical, intradermal, intranasal or intrabronchial administration. The dosage regimen will be determined by the attending physician and clinical factors. As is well known in the medical arts, dosages for any one patient depends upon many factors, including the patient's size, body surface area, age, the particular compound to be administered, sex, time and route of administration, general health, and other drugs being administered concurrently. A typical dose can be, for example, in the range of 0.001 to 1000 µg (or of nucleic acid for expression or for inhibition of expression in this range); however, doses below or above this exemplary range are envisioned, especially considering the aforementioned factors. Generally, the regimen as a regular administration of the pharmaceutical composition should be in the range of 1 µg to 10 mg units per day. If the regimen is a continuous infusion, it should also

be in the range of 1 µg to 10 mg units per kilogram of body weight per minute, respectively. Progress can be monitored by periodic assessment. Dosages will vary but a preferred dosage for intravenous administration of DNA is from approximately 10^6 to 10^{12} copies of the DNA molecule. The compositions of the invention may be administered locally or systemically. Administration will generally be parenterally, e.g., intravenously; DNA may also be administered directly to the target site, e.g., by biolistic delivery to an internal or external target site or by catheter to a site in an artery. Preparations for parenteral administration include sterile aqueous or non-aqueous solutions, suspensions, and emulsions. Examples of non-aqueous solvents are propylene glycol, polyethylene glycol, vegetable oils such as olive oil, and injectable organic esters such as ethyl oleate. Aqueous carriers include water, alcoholic/aqueous solutions, emulsions or suspensions, including saline and buffered media. Parenteral vehicles include sodium chloride solution, Ringer's dextrose, dextrose and sodium chloride, lactated Ringer's, or fixed oils. Intravenous vehicles include fluid and nutrient replenishers, electrolyte replenishers (such as those based on Ringer's dextrose), and the like. Preservatives and other additives may also be present such as, for example, antimicrobials, anti-oxidants, chelating agents, and inert gases and the like. Furthermore, the pharmaceutical composition of the invention may comprise further agents such as interleukins or interferons depending on the exact intended use of the pharmaceutical composition.
The above methods referring to downstream developments also apply to therapeutically effective compounds referred to in additional embodiments herein below.

In another embodiment of the invention a method is disclosed of selecting a composition for inhibiting leukemia in a patient, the method comprising separately maintaining aliquots of cells of a patient sample in the presence of a plurality of test compositions, comparing the expression profile, preferably the level of expression of a group of markers in each of the aliquots, and selecting one of the test compositions which induces an altered expression profile of the group of markers in the aliquot containing that test composition, relative to other test compositions characterized in that the group of markers consists of markers

selected independently from the markers listed in one or more of the tables 1 to 20, tables 25 or 27 or tables 29, 30, 32, 33, 35, 36, 38, 39, 41, 42 and whereby the number of markers in the group is between one, preferably two such as 3, 4, 5, 6, 7, 8, 9 or 10 and the total number of markers listed in the tables 1 to 20, tables 25 or 27 or tables 29, 30, 32, 33, 35, 36, 38, 39, 41, 42.

Again, as with the previously recited embodiments, the alteration determined in accordance with the method of the invention in the expression profile or expression level must be in the direction of the expression profile of normal cells or at least diseased but non-leukemic cells. Accordingly, it is also preferred in accordance with this embodiment that the comparison includes an internal standarad of expression levels of analysed markers wherein the internal standarad represents the expression profile of non-leukemic and preferably normal cells. The comparison may – again – be effected by relying on actual experimental data or on in silico obtained reference data.

The expression "in the direction of the expression profile of normal cells" as used herein preferably relates to cells that comprise blood cells, more preferably a single type of blood cells. Most preferably, the single type of cells corresponds to the type of the leukemic cell. For example, an AML type of leukemic cell would preferably be compared to a healthy myeloic blast cell whereas a ALL type of leukemic cell would preferably be compared to a healthy lymphatic blast cell. Myeloic blast cells and lymphatic blast cells may be isolated from healthy bone marrow using well known methods, such as cell sorting based on flow cytometry using established cell surface markers.

In this method of the invention, it is preferred that the test composition comprises only one putatively active test compound. Insofar, the correlation with the activity of the test compound and the readout is particularly convenient. If the test composition comprises more than one putatively pharmaceutically active compounds, it may be considered to separately test each compound in a composition that has tested positive in a first round of the assay. Consequently, in a second round, i.e. in a repetition of steps (a) and (b), the various compositions

tested positive, if any, in the first round, may be subdivided into single compounds and these single compounds tested again for their efficacy. The goal of such an approach, of course, is to obtain a composition comprising a single active compound only.

In another embodiment a method of determining new subtypes of leukemia cells is diclosed, the method comprising determining. the expression profile, preferably the level of expression of a group of markers of leukemia cells of unknown subtype, comparing the expression profile to the level of expression, ie. the expression profile, of a group of markers of leukemia cells of known subtype, thereby concluding that a new subtype is determined when the expression profile, preferably the level of expression is different to all known subtypes characterized in that the group of markers consists of markers selected independently from the markers listed in one or more of the tables 1 to 20, tables 25 or 27 or tables 29, 30, 32, 33, 35, 36, 38, 39, 41, 42 and whereby the number of markers in the group is between one, preferably two and the total number of markers listed in the tables 1 to 20, tables 25 or 27 or tables 29, 30, 32, 33, 35, 36, 38, 39, 41, 42.

The term "subtype of leukemia cells" in accordance with the present invention may be better understood in accordance with the following Leukemias are subdivided according to their natural clinical course into acute and chronic leukemias. Based on the cell line they are derived from they are further subdivided into myeloid and lymphatc leukemias. This results in four leukemia types, i.e. acute myeloid leukemia (AML), acute lymphoblastic leukemia (ALL), chronic myeloid leukemia (CML), and chronic lymphatic leukemia (CLL). Based on genetic, phenotypic, and biological characteristic, which are assessed by cytomorphology, cytochemistry, cytogenetics, immunophenotyping, and molecular genetics, AML, ALL, and CLL are further subdivided into subtypes. These subtypes are associated with highly differing prognoses. Treatment approaches specific for these subtypes are applied and are being further optimized. Thus, an exact diagnosis based on a reliable and reproducible method is essential for the selection of the appropriate subtype-specific treatment.

The new subtypes identified in accordance with the invention may then be subjected in the same or in further patients to the other methods/embodiments of the invention.

In another embodiment a method is disclosed for guiding the therapy of leukemia in a patient depending on the leukemia subtype and/or the risk of relapse of disease, the method comprising determining the expression profile, preferably the level of expression of a group of markers in the patient sample, and deciding about the therapy strategy depending on the leukemia subtype or the risk of relapse of disease characterized in that the group of markers consists of markers selected independently from the markers listed in one or more of the 1 to 20, tables 25 or 27 or tables 29, 30, 32, 33, 35, 36, 38, 39, 41, 42 and whereby the number of markers in the group is between one, preferably two such as 3, 4, 5, 6, 7, 8, 9 or 10 and the total number of markers listed in the tables 1 to 20, tables 25 or 27 or tables 29, 30, 32, 33, 35, 36, 38, 39, 41, 42.

This embodiment is particularly important for the quick and reliable recovery of the patient from the leukemia that effects him or her. As has been stated above, the early and reliable diagnosis of the leukaemia type or subtype is particularly important for the instigation of a useful and straightforward treatment regimen. An incorrect diagnosis may result in the application of a wrong treatment regimen which, in turn, may lead to significant health risks including premature death of the patient. In accordance with the present invention, a reliable means has been provided that, based on the inventive selection of markers provided, will overcome the prior art problems of an insecure or an inappropriate time frame demanding diagnosis. In particular, the present method of the invention provides in step (a) an unambiguous and safe basis for the decision step (b). Again, the patient may safely rely on the conclusion drawn in step (b) due to the strong inherent correlation that has been achieved between the selection of markers and the leukemia subtype. The relation of tables to leukemia subtypes has also been demonstrated elsewhere in this specification.

In another embodiment of the invention, a method for monitoring the progression of leukemia in a patient is disclosed, the method comprising determining the the expression profile, preferably the level of expression of a group of markers in a patient sample at a first point in time, and repeating this step at a subsequent point in time; and comparing the expression profile, preferably the level of expression detected in the previous steps and therefrom monitoring the progression of leukemia in the patient, characterized In that the group of markers consists of markers selected independently from the markers listed in one or more of the tables 1 to 20, tables 25 or 27 or tables 29, 30, 32, 33, 35, 36, 38, 39, 41, 42 and whereby the number of markers in the group is between one, preferably two such as 3, 4, 5, 6, 7, 8,.9 or 10 and the total number of markers listed in the tables 1 to 20, tables 25 or 27 or tables 29, 30, 32, 33, 35, 36, 38, 39, 41, 42. In a preferred embodiment, the patient has undergone chemotherapy between the first point in time and the subsequent point in time (including repetitions of step (b).

In this embodiment of the present invention, the skilled artisan may repeat step (b) one or more times in order to collect additional data from different (more) time points. The additional data obtained by such further measurements may provide an overall better overview on the progress of the disease.
In accordance with this embodiment of the disease, the term "progression of leukemia" includes the interpretation of "regression of leukemia", i.e. includes the interpretation of a negative progression. This is of course in line with the aim of the therapy and the desire of the patient.

In the methods according to the invention it is preferred that the group of markers consists of markers selected independently from the markers listed in one or more of the tables 1 to 20, tables 25 or 27 or tables 29, 30, 32, 33, 35, 36, 38, 39, 41, 42 and whereby the number of markers in the group is between one, preferably two and the total number of markers listed in the at least one of tables 1 to 20, tables 25 or 27 or tables 29, 30, 32, 33, 35, 36, 38, 39, 41, 42. In a preferred embodiment, the number of markers in the group is between five, more preferably between 7, 10 or 15 and the total number of markers listed in the tables 1 to 20, tables 25 or 27 or tables 29, 30, 32, 33, 35, 36, 38, 39, 41, 42. It is feasible that the

group of markers not only consists of those markers but also comprises them as the data will then be still statistically significant, i.e. the preferred groups may additionally contain 10, 50 or 100 other markers and comprise the other markers according to the invention and mentioned above. It is, however, also feasible for the expert skilled in the art that only a single suitable marker is determined with the methods according to the invention.

Particulary preferred markers used in a method where only one or a few as e.g. one, preferably two markers are used are described in Table 22 and Example 3, Fig. 12 or the markers marked with an asterisk in table 20 and shown in tables 16 to 19 as the preferred set of markers. In detail, example 3 mentions (see example 3 for more details) the following markers including their expression level:

- *ADCY3*
- adenosine deaminase (*ADA*)
- *ARGHGAP4*
- B-cell specific coactivator of octamer binding transcription factors
- *CAPN3* is a member of the papain superfamily and was higher expressed in CML
- *CBFß-MYH11*
- *CD24*
- *CD27*, was identified to assign samples either ALL or CLL
- *CD74* plays a critical role in MHC class II antigen processing
- connective tissue growth factor (*CTGF*)
- *CTGF*
- *CTSW*
- *MYH11*
- glucocorticoid receptor beta
- higher expression of *CBFA2T1* (formerly *ETO*)
- *HLA-DMB*
- *HOXA9*
- *HOXB5*
- *IRF4*, an immune system-restricted interferon regulatory factor
- *KIAA1013*

- *LCN2* that shown to be a modulator of inflammation
- *LEF-1* was absent in myeloid leukemias but highly expressed in lymphoid leukemias
- *MBNL*
- *MSF* translocation partner of the mixed-lineage leukemia gene (MLL) in AML
- *NCOA1* expressed higher in CLL as compared to ALL
- *OS-9*differentially expressed between AML and ALL (14)
- Phospholipidscramblase 1 (*PLSCR1*) to be lower expressed in AML and ALL as compared to normal BM
- *POU2AF1*
- *POU2F2*
- *POU4F1*
- *SCYA3*
- *SGP28*
- *SOCS-2*
- *TRB and CD3D*

Particulary preferred markers used in a method where only one or a few as e.g. one, preferably two markers are used are described in tables 30, 33, 36 and 42 and Example 7, Figures 189 to 234, 254 to 272, 338 to 371, 433 to 465, respectively, or the markers marked with an asterisk in tables 29, 32, 35, 38, and 41 and Figures 24 to 188, 235 to 253, 273 to 337, 372 to 405, 406 to 432, respectively as the preferred set of markers. In detail, example 7 mentions (see example 7 for more details) the following markers including their expression level:

geneID	gene symbol	feature
201162_at	IGFBP7	CLL low
201163_s_at	IGFBP7	CLL low
201362_at	NS1-BP	CML high

201496_x_at	MYH11	AML inv(16) high
201497_x_at	MYH11	AML inv(16) high
201998_at	SIAT1	CLL high
202095_s_at	BIRC5	CLL low
203074_at	ANXA8	AML t(15;17) high
204150_at	STAB1	AML t(15;17) high
204511_at	KIAA0793	CLL high
205528_s_at	CBFA2T1	AML t(8;21) high
205529_s_at	CBFA2T1	AML t(8;21) high
205805_s_at	ROR1	CLL high
206940_s_at	POU4F1	AML t(8;21) high
207819_s_at	ABCB4	CLL high
208091_s_at	DKFZP564K0822	CLL high
208456_s_at	RRAS2	CLL high
209061_at	NCOA3	CLL high
209101_at	CTGF	ALL t(4;11) high, ALL Ph high, T-ALL high
209374_s_at	IGHM	CLL high
209616_s_at	CES1	AML MLL high

210997_at	HGF	AML t(15;17) high
212285_s_at	AGRN	AML t(15;17) high
213539_at	CD3D	T-ALL high
214450_at	CTSW	AML t(15;17) high
215925_s_at		ALL t(4;11) high
218223_s_at	LOC51177	CML low
222166_at		AML +8 high
224520_s_at	MGC13168	ALL t(8;14) high
224794_s_at	LOC51148	AML t(15;17) high
225660_at	SEMA6A	ALL B not Ph high, ALL Ph high
226496_at	Homo sapiens, Similar to hypothetical protein FLJ22611, clone MGC:24716 IMAGE:4277726, mRNA, complete cds	ALL high, CLL high
228827_at	Homo sapiens clone 25023 mRNA sequence	AML t(8;21) high

228904_at	ESTs	AML normal high, AML +8 high, AML complex high
236301_at	Homo sapiens, clone IMAGE:3866403, mRNA	CLL high
236892_s_at	HOXB6	AML normal high, AML +8 high, AML complex high
239214_at	ESTs	ALL t(4;11) high
239393_at	ESTs	ALL t(4;11) high
239791_at	HOXB6	AML normal high, AML +8 high
240581_at	ESTs	ALL t(4;11) high
241464_s_at	ESTs	AML MLL high, AML normal high, AML +8 high, AML complex high
241525_at	ESTs	AML inv(16) high
243362_s_at	LEF1	ALL high, CLL high
36566_at	CTNS	T-ALL low
38487_at	FLJ12442	AML t(15;17) high
201105_at	LGALS1	ALL t(4;11) high

204044_at	QPRT	ALL t(4;11) high
205899_at	CCNA1	ALL t(4;11) high
209168_at	GPM6B	ALL t(4;11) high
213539_at	CD3D	T-ALL high
213894_at	KIAA0960	ALL t(4;11) high
215925_s_at		ALL t(4;11) high
218224_at	PNMA1	T-ALL high
219463_at	C20orf103	ALL t(4;11) high
219631_at	FLJ12929	T-ALL high
225563_at	ESTs	ALL t(4;11) high
225592_at	NRM	ALL t(4;11) high
228083_at	Homo sapiens mRNA; cDNA DKFZp434I1216 (from clone DKFZp434I1216)	ALL t(4;11) high
228988_at	ZNF6	T-ALL high
235749_at		ALL t(8;14) high
242414_at	ESTs	ALL t(4;11) high
243756_at	ESTs	ALL t(4;11) high

201497_x_at	MYH11	AML inv(16) high
228827_at	Homo sapiens clone 25023 mRNA sequence	AML t(8;21) high
38487_at	FLJ12442	AML t(15;17) high
203074_at	ANXA8	AML t(15;17) high
205528_s_at	CBFA2T1	AML t(8;21) high
205529_s_at	CBFA2T1	AML t(8;21) high
206940_s_at	POU4F1	AML t(8;21) high
211341_at	POU4F1	AML t(8;21) high
201496_x_at	MYH11	AML inv(16) high
228660_x_at	SEMA4F	other high
202718_at	IGFBP2	AML t(15;17) high
205380_at	PDZK1	other high
202746_at		AML MLL low
201596_x_at	KRT18	AML t(8;21) low
34210_at	CDW52	AML t(15;17) low
212850_s_at	LRP4	AML inv(16) high

228904_at	ESTs	AML t(8;21) low, AML t(15;17) low, AML inv(16) low, AML MLL low
203151_at	MAP1A	AML t(8;21) low
201137_s_at	HLA-DPB1	AML t(15;17) low
200675_at	CD81	AML inv(16) low
201425_at	ALDH2	AML t(8;21) low
202085_at	TJP2	AML inv(16) low
202619_s_at	PLOD2	AML MLL low
203092_at	TIMM44	AML inv(16) low
204425_at	ARHGAP4	AML t(15;17) low
205366_s_at	HOXB6	AML t(8;21) low, AML t(15;17) low, AML inv(16) low, AML MLL low
205472_s_at	DACH	AML MLL high
206761_at	TACTILE	AML MLL low
222166_at		AML +8 low
222335_at	ESTs	AML MLL low
223318_s_at	MGC10974	AML complex low

225330_at	Homo sapiens, clone MGC:18216 IMAGE:4156235, mRNA, complete cds	AML inv(16) low
231277_x_at	ESTs	AML complex low
635_s_at	PPP2R5B	other low
202503_s_at	KIAA0101	CLL low
202580_x_at	FOXM1	CLL low
202709_at	FMOD	CLL high
204882_at	KIAA0053	CLL high
205049_s_at	CD79A	ALL high, CLL high
205051_s_at	KIT	AML high
205382_s_at	DF	AML high
205599_at	TRAF1	CML low CLL high
206255_at	BLK	ALL high, CLL high
206398_s_at	CD19	ALL high, CLL high
210487_at	DNTT	ALL high
210948_s_at	LEF1	ALL high, CLL high
211352_s_at	NCOA3	CLL high

211404_s_at	APLP2	AML high
214761_at	OAZ	ALL high
217950_at	NOSIP	CLL high
218090_s_at		CLL high
218516_s_at	FLJ20421	normal BM low
218916_at	FLJ23436	normal BM low
219753_at	STAG3	ALL high
221969_at	PAX5	ALL high, CLL high
223703_at	CDA017	AML high, CML high, normal BM high
226147_s_at	Homo sapiens cDNA: FLJ22667 fis, clone HSI08385	CLL high
228471_at	ESTs	CLL high
229487_at	ESTs	ALL high
229790_at	TERF2	CML low, BM low
231736_x_at	MGST1	AML high, CML high, normal BM high
231854_at	Homo sapiens cDNA FLJ11448 fis, clone HEMBA1001391	CML low

239287_at	ESTs	CLL high
243362_s_at	LEF1	ALL high
243363_at	LEF1	ALL high, CLL high
41577_at	PPP1R16B	CML low

Preferred methods for detection and quantification of the amount of nucleic acids, i.e. for the methods according to the invention allowing the determination of the level of expression of a marker or a group of markers, are those described by Sambrook et al. (1989) or real time methods known in the art as the TaqMan® method disclosed in WO92/02638 and the corresponding US patents US 5,210,015, US 5,804,375, US 5,487,972. This method exploits the exonuclease activity of a polymerase to generate a signal. In detail, the (at least one) target nucleic acid component is detected by a process comprising contacting the sample with an oligonucleotide containing a sequence complementary to a region of the target nucleic acid component and a labeled oligonucleotide containing a sequence complementary to a second region of the same target nucleic acid component sequence strand, but not including the nucleic acid sequence defined by the first oligonucleotide, to create a mixture of duplexes during hybridization conditions, wherein the duplexes comprise the target nucleic acid annealed to the first oligonucleotide and to the labeled oligonucleotide such that the 3'-end of the first oligonucleotide is adjacent to the 5'-end of the labeled oligonucleotide. Then this mixture is treated with a template-dependent nucleic acid polymerase having a 5' to 3' nuclease activity under conditions sufficient to permit the 5' to 3' nuclease activity of the polymerase to cleave the annealed, labeled oligonucleotide and release labeled fragments. The signal generated by the hydrolysis of the labeled oligonucleotide is detected and/ or measured. TaqMan® technology eliminates the need for a solid phase bound reaction complex to be formed and made detectable. Other methods include e.g. fluorescence resonance energy transfer between two adjacenly hybridized probes as used in the LightCycler® format described in US 6,174,670.

Protocols for carrying out the methods according to the invention are known to the expert in the field and are described in the examples, preferably in example 1 and 4. A preferred protocol is described in Example 1(A), where total RNA is isolated, cDNA synthesized and biotin incorporated during the transcription reaction. The purified cDNA was applied to commercially available arrays which can be obtained e.g. from Affymetrix. The hybridized cDNA is detected according to the methods described in Example 1(A). The arrays are produced by photolithography or other methods known to experts skilled in the art e.g. from US5,445,934, US5,744,305, US5,700,637, US5,945,334 and EP619 321 or EP 373 203. The latter methods are also suitable for producing the composition according to the inventions in particular the composition wherein polynucleotides or oligonucleotides are bound to a solid phase in particular in the form of arrays. In a further preferred embodiment of the methods according to the invention, a transcribed polynucleotide or portion thereof is the marker or at least one of the markers. A particularly preferred transcribed polynucleotide is an mRNA or a cDNA. In a preferred embodiment of the methods according to the invention, the step of determining the expression profile further comprises amplifying the transcribed polynucleotide. In another preferred embodiment, the level of expression, i.e. the expression profile, of the group of transcribed polynucleotides is determined by annealing the transcribed polynucleotides with a complementary polynucleotide or a portion thereof under stringent hybridization conditions. The term "stringent hyberidisation conditions" is equivalent to the term "highly stringent hyberdisation conditions". Such highly stringent hybridization conditions may be determined in accordance with the teachings provided in Hames and Higgins (eds) "Nucleic acid hybridization, a practical approach", IRL Press 1985, Oxford, and include hybridization at 55-65°C in 0.2-0.5xSSC, 0.1% SDS followed by appropriate washing conditions such as 0.5-1xSSC at 55°C and 0.1% SDS.

In a most preferred embodiment, the patient sample is blood, i.e. blood mononuclear cells, or bone marrow, i.e. mononuclear cells. The methods according to the invention may be performed on fresh or frozen blood, i.e. blood mononuclear cells or bone marrow, i.e. mononuclear cells.

In a preferred embodiment the marker or at least one of the markers is a protein. In another preferred embodiment the expression profile of the proteins is detected using a reagent which specifically binds to one of the proteins whereby preferably the reagent is selected from the group consisting of an antibody, an antibody derivative, and an antibody fragment.

The term "antibody" comprises monoclonal antibodies as first described by Köhler and Milstein in Nature 278 (1975), 495-497 as well as polyclonal antibodies, i.e. entibodies contained in a polyclonal antiserum. Monoclonal antibodies include those produced by transgenic mice. Fragments of antibodies include $F(ab')_2$, Fab and Fv fragments. Derivatives of antibodies include scFvs, chimeric and humanized antibodies. See, for example Harlow and Lane, loc. cit.

Another embodiment of the invention is a kit preferably for assessing the suitability of each of a plurality of compounds for inhibiting leukemia in a patient, the kit optionally comprising the plurality of compounds; and a reagent for assessing the expression profile of a group of markers characterized in that the group of markers consists of markers selected independently from the markers listed in one or more of the tables 1 to 20, tables 25 or 27 or tables 29, 30, 32, 33, 35, 36, 38, 39, 41, 42 and whereby the number of markers in the group is between two and the total number of markers listed in the tables 1 to 20, tables 25 or 27 or tables 29, 30, 32, 33, 35, 36, 38, 39, 41, 42. Another embodiment is a kit preferably for assessing whether a patient is afflicted with leukemia, the kit comprising reagents for assessing the expression profile of a group of markers characterized In that the group of markers consists of markers selected independently from the markers listed in one or more of the tables 1 to 20, tables 25 or 27 or tables 29, 30, 32, 33, 35, 36, 38, 39, 41, 42 and whereby the number of markers in the group is between two and the total number of markers listed in the tables 1 to 20, tables 25 or 27 or tables 29, 30, 32, 33, 35, 36, 38, 39, 41, 42. Another embodiment is a kit preferably for assessing the presence of human leukemia cells, the kit comprising an antibody, wherein the antibody specifically binds with a protein corresponding to a marker characterized In that the marker is selected from the tables 1 to 20, tables 25 or 27 or tables 29, 30, 32, 33, 35, 36, 38, 39, 41, 42. Another

embodiment is a kit preferably for assessing the leukemia cell carcinogenic potential of a test compound, the kit comprising leukemia cells and a reagent for assessing expression of a marker, wherein the marker is selected from the tables 1 to 20, tables 25 or 27 or tables 29, 30, 32, 33, 35, 36, 38, 39, 41, 42.

Advantageously, the kit of the present invention further comprises, optionally (a) storage solution(s) and/or remaining reagents or materials required for the conduct of scientific and/or diagnostic and/or therapeutic methods. Furthermore, parts of the kit of the invention can be packaged individually in vials or bottles or in combination in containers or multicontainer units.

Another embodiment of the invention is related to a protein or the RNA, cDNA or cRNA corresponding to a marker selected from the tables 1 to 20, tables 25 or 27 or tables 29, 30, 32, 33, 35, 36, 38, 39, 41, 42 or the use thereof for the treatment of or vaccination against leukemia. Alternatively and depending on the exact purpose, inhibitors of these compounds such as antibodies, fragments or derivatives thereof may be employed for said purpose.

The invention also contemplates a method for the development or preparation of a pharmaceutical composition for the treatment of leukemia characterized in that a protein corresponding to a marker selected from the tables 1 to 20, tables 25 or 27 or tables 29, 30, 32, 33, 35, 36, 38, 39, 41, 42 is admixed with pharmaceutical compounds. Another embodiment of the invention is related to a method for the development or preparation of a pharmaceutical composition for the treatment of leukemia characterized in that a vector comprising a polynucleotide encoding a protein corresponding to a marker selected from the tables 1 to 20, tables 25 or 27 or tables 29, 30, 32, 33, 35, 36, 38, 39, 41, 42 is admixed with pharmaceutical compounds. Another embodiment of the invention is a method for the development or preparation of a pharmaceutical composition for the treatment of leukemia characterized in that an antisense oligonucleotide complementary to a polynucleotide encoding a protein corresponding to a marker selected from the tables 1 to 20, tables 25 or 27 or tables 29, 30, 32, 33, 35, 36, 38, 39, 41, 42 is admixed with pharmaceutical compounds. Alternatively, inhibitors such as

antibodies specific for the markers may be used for the preparation or development of a pharmaceutical composition.

The term "pharmaceutical compounds" is preferably to be understood to mean pharmaceutically acceptable carriers, diluents or excipients, only in connection with the embodiments recited in this paragraph. In another embodiment of the invention a marker or a group of markers selected individually from one or more of the tables 1 to 20, tables 25 or 27 or tables 29, 30, 32, 33, 35, 36, 38, 39, 41, 42 is used for the determination of leukemia cells, the type or subtype of leukemia cells.

In another embodiment of the invention a marker or a group of markers selected individually from one or more of the tables 1, 2, 13, 14, 17, 25, 27, 35 or 36 is used for the determination of the subtype of AML cells.

In a preferred embodiment, the invention is related to a composition comprising a group of markers and substances chemically different to the markers characterized in that the group of markers consists of markers selected independently from the markers listed in one or more of the tables 1 to 20, tables 25 or 27 or tables 29, 30, 32, 33, 35, 36, 38, 39, 41, 42 and whereby the number of markers in the group is between one, preferably two and the total number of markers listed in the tables 1 to 20, tables 25 or 27 or tables 29, 30, 32, 33, 35, 36, 38, 39, 41, 42. It is preferred that the composition according to the invention is characterized in that the group of markers consists of all markers listed in one or more of the tables 1 to 20, tables 25 or 27 or tables 29, 30, 32, 33, 35, 36, 38, 39, 41, 42. More preferred the composition according to the invention is characterized in that the group of markers consists of all markers listed in one or more of the tables 14, tables 16 to 20, or table 29 or 30, most preferred the group of markers consists of all markers listed in the tables 16 to 20 or tables 29 or 30. Preferably the markers are polynucletides or oligonucleotides, whereby the polynucleotides are bound to a solid phase in the form of an array.

The present invention also relates to a method of determining the subtypes of ALL cells in a patient sample comprising the steps of a) determining the level of expression of a group of markers in the patient sample and b) concluding from the differences in the level of expression which subtypes of ALL cells the patient sample contains characterized in that the group of markers consists of markers selected independently from the markers listed in one or more of the tables 18, 32 or 33 and whereby the number of markers in the group is between two and the total number of markers listed in the tables 18, 32 or 33.

Preferably the group of markers consists of all markers listed in one or more of the tables 18, 32 or 33.

The present invention further relates to a method of determining the subtypes of CLL cells in a patient sample comprising the steps of a) determining the level of expression of a group of markers in the patient sample and b) concluding from the differences in the level of expression which subtypes of CLL cells the patient sample contains characterized in that the group of markers consists of markers selected independently from the markers listed in one or more of the tables 38 or 39 and whereby the number of markers in the group is between two and the total number of markers listed in the tables 38 or 39.

It is preferred that the group of markers consists of all markers listed in one or more of the tables 38 or 39.

The present invention is also related to a diagnostic composition comprising at least one nucleic acid molecule, preferably (a) single-stranded nucleic acid molecule(s), which is capable of specifically hybridizing to the mRNA of at least one gene listed in Table 1. The use of said nucleic acid molecules for diagnosis of leukemia subtypes, preferably based on microarray technology, offers the following advantages: (1) more rapid and more precise diagnosis, (2) easy to use in laboratories without specialized experience, (3) abolishes the requirement for analyzing viable cells for chromosome analysis (transport problem), (4) very experienced hematologists for cytomorphology and cytochemistry, immunophenotyping as well as cytogeneticists and molecularbiologists are no longer required, and (5) improves the subclassification of leukemia due to the

definition of new entities based on gene expression profiles in those subtypes that are not clearly defined with the methods of the prior art (class discovery).

As used herein, the term "capable of specifically hybridizing" has the meaning of hybridization under conventional hybridization conditions, preferably under stringent conditions as described, for example, in Sambrook, J., et al., in "Molecular Cloning: A Laboratory Manual" (1989), Eds. J. Sambrook, E. F. Fritsch and T. Maniatis, Cold Spring Harbour Laboratory Press, Cold Spring Harbour, NY and the further definitions provided above. Also contemplated are nucleic acid molecules that hybridize at lower stringency hybridization conditions. Changes in the stringency of hybridization and signal detection are primarily accomplished through the manipulation, preferably of formamide concentration (lower percentages of formamide result in lowered stringency), salt conditions, or temperature. For example, lower stringency conditions include an overnight incubation at 37°C in a solution comprising 6X SSPE (20X SSPE = 3M NaCl; 0.2M NaH2PO4; 0.02M EDTA, pH 7.4), 0.5% SDS, 30% formamide, 100 mg/ml salmon sperm blocking DNA, followed by washes at 50°C with 1 X SSPE, 0.1% SDS. In addition, to achieve even lower stringency, washes performed following stringent hybridization can be done at higher salt concentrations (e.g. 5x SSC). Variations in the above conditions may be accomplished through the inclusion and/or substitution of alternate blocking reagents used to suppress background in hybridization experiments. The inclusion of specific blocking reagents may require modification of the hybridization conditions described above, due to problems with compatibility.

As a hybridization probe (or primer) nucleic acid molecules can be used, for example, that have exactly or basically the nucleotide sequence of at least one of the genes depicted in the appended tables or parts of these sequences. The term nucleic acid molecule as used herein also comprises fragments which are understood to be parts of the nucleic acid molecules that are long enough to specifically hybridize to transcripts of at least one of the genes of the appended tables. These nucleic acid molecules can be used, for example, as probes or primers in a diagnostic assay. Preferably, the nucleic acid molecules of the present invention have a length of at least 8, 10, 12, 13, 15, 18 in particular of at least 20 and particular preferred of at least 25 nucleotides. The nucleic acid molecules of the invention or parts therefrom* can also be used, for example, as primers for a PCR reaction. The fragments used as hybridization probe can be

synthetic fragments that were produced by means of conventional synthesis methods.

In a preferred embodiment, the diagnostic composition of the present invention comprises at least nucleic acid molecules which are capable of specifically hybridizing to the mRNAs of at least one of the genes listed in the appended tables, preferably 2-5, more preferably 8-12 genes.

In a more preferred embodiment, the diagnostic composition of the present invention comprises at least nucleic acid molecules which are capable of specifically hybridizing to the mRNAs of at least one of the genes listed in the appended tables. In a further preferred embodiment, the diagnostic composition of the present invention comprises at least nucleic acid molecules which are capable of specifically hybridizing to the mRNAs of all genes listed in the appended tables.

In a further preferred embodiment, the nucleic acid molecules of the diagnostic composition of the present invention are bound to (a) a solid support, for example, a polystyrene microtiter dish or nitrocellulose membrane or glass surface or (b) to non-immobilized particles in solution.

In an even more preferred embodiment, the nucleic acid molecules of the diagnostic composition are present in a microarray format which can be established according to well known methods; for details see, e.g., www.affymetrix.com/technology/tech_spotted.html; www.affymetrix.com/technology/tech_probe.html.

The present invention also provides the use of (a) nucleic acid molecule(s) of the present invention for the preparation of a diagnostic composition for the diagnosis of a leukemia or for the diagnosis of several subtypes or a disposition to a leukemia. For the diagnosis of a particular leukemia subtype, preferably, at least 5 different nucleic acid molecules are used as probes. For diagnosis, preferably, bone marrow or peripheral blood can be used. For diagnosis, the target sample is contacted with a (a) nucleic acid molecule(s) of the present invention and the concentration of individual mRNAs is compared with the mRNA expression profile levels of a test sample obtained from healthy donors.

It is a further embodiment of the invention to provide a method of determining whether a patient sample contains leukemia cells or other cells and at the same

time determining the type and subtype of leukemia cells, comprising the steps of providing a patient sample, isolating RNA from the patient sample, transcribing the RNA into cDNA and transcribing the cDNA into cRNA while simultaneously labelling the cRNA, hybridising the cRNA to a microarray, and determining the level of expression of a marker or a group of markers.

Further, the invention contemplates the use of a marker or a group of markers for determining whether a patient sample contains leukemia cells or other cells and whereby preferably the type and subtype of leukemia cells is simultaneously or subsequently is determined. The markers specified in the appended examples and tables may, in accordance with the invention, be used to differentiate, for example, between ALL, CLL, CML and AML.

The nucleic acid molecule is typically a nucleic acid probe for hybridization or a primer for PCR. The person skilled in the art is in a position to design suitable nucleic acids probes based on the information provided in in the appended tables.

The target cellular component, i.e. mRNA e.g., in bone marrow or blood (BM) may be detected directly in situ, e.g. by in situ hybridization or it may be isolated from other cell components by common methods known to those skilled in the art before contacting with a probe. Detection' methods include Northern blot analysis, RNase protection, in situ methods, e.g. in situ hybridization, in vitro amplification methods (PCR, LCR, QRNA replicase or RNA-transcription/amplification (TAS, 3SR), reverse dot blot disclosed in EP 0 237 362)) and other detection assays that are known to those skilled in the art. Preferably, detection is based on a microarray.

Amplification methods include the polymerase chain reaction (PCR) which specifically amplifies target sequences to detectable amounts. Other possible amplification reactions are the ligase Chain Reaction (LCR, Wu and Wallace, 1989, Genomics 4:560-569 and Barany, 1991, Proc. Natl. Acad. Sci. USA 88:189-193); Polymerase Ligase Chain Reaction (Barany, 1991, PCR Methods and Applic. 1:5-16); Gap-LCR (PCT Patent Publication No. WO 90/01069); Repair Chain Reaction (European Patent Publication No. 439,182 A2), 3SR (Kwoh et al., 1989, Proc. Natl. Acad. Sci. USA 86:1173-1177; Guatelli et al., 1990, Proc. Natl. Acad. Sci. USA 87:1874-1878; PCT Patent Publication No. WO 92/0880A), and NASBA (U.S. Pat. No. 5,130,238). Further, there are strand displacement amplification (SDA), transciption mediated amplification (TMA), and Q□-

amplification (for a review see e.g. Whelen and Persing (1996). Annu. Rev. Microbiol. 50, 349-373; Abramson and Myers, 1993, Current Opinion in Biotechnology 4:41-47).

Products obtained by in vitro amplification can be detected according to established methods, e.g. by separating the products on agarose gels and by subsequent staining with ethidium bromide. Alternatively, the amplified products can be detected by using labeled primers for amplification or labeled dNTPs.

The probes can be detectably labeled, for example, with a radioisotope, a bioluminescent compound, a chemiluminescent compound, a fluorescent compound, a metal chelate, biotin or an enzyme.

The invention further contemplates a method of making an isolated hybridoma which produces an antibody useful for assessing whether a patient is afflicted with leukemia, the method comprising isolating a protein corresponding to a marker selected from the group consisting of the markers listed in Tables 1 to 20, tables 25 or 27 or tables 29, 30, 32, 33, 35, 36, 38, 39, 41, 42 immunizing a mammal using the isolated protein, or a peptide corresponding to its sequence or a part thereof; isolating splenocytes from the immunized mammal-, fusing the isolated splenocytes with an immortalized cell line to form hybridomas; and screening individual hybridomas for production of an antibody which specifically binds with the protein to isolate the hybridoma. Further, an antibody produced by this method is contemplated by the invention. The antibody may be fragmented or derivated to obtained fragment or derivatives retaining the antibody specificity as has been described herein above.

The invention further contemplates a method of assessing the leukemia cell carcinogenic potential of a test compound, the method comprising maintaining separate aliquots of leukemia cells in the presence and absence of the test compound; and comparing expression of a marker in each of the aliquots, wherein a significantly altered level of expression of the marker in the aliquot maintained in the presence of the test compound, relative to the aliquot maintained in the absence of the test compound, is an indication that the test compound possesses human breast cell carcinogenic potential wherein a marker according to the invention is used.

The invention further contemplates a system for identifying selected polynucleotide records that identify a leukemia cell, the system comprising: a digital computer-, a database coupled to the computer; a database coupled to the database server having data stored in, the data comprising records of data comprising a polynucleotide corresponding to a marker according to the invention and a code mechanism for applying queries based upon a desired selection criteria to the data file in the database to produce reports of polynucleotide records which match the desired selection criteria.

The invention also relates to a method for detecting a leukemia cell, using a computer having a processor, memory, display, and input/output devices, the method comprising the steps of

a) providing a sequence of a polynucleotide isolated from a sample suspected of containing a leukemia cell,

b) providing a database comprising records of data comprising a polynucleotide corresponding to a group of markers according to the invention;

c) using a code mechanism for applying queries based upon a desired selection criteria to the data file in the database to produce reports of polynucleotide records of step a) which provide a match of the desired selection criteria of the sequences in the database of step b), the presence of a match being a positive indication that the polynucleotide of step 1) has been isolated from a cell that is a-leukemia cell.

Also, the present invention relates to a method for assessing the leukemia cell carcinogenic potential of a test compound, comprising (a) contacting a non-leukemia cell with a test compound, and (b) assessing an increase or decrease of marker expression in said non-leukemia cell wherein the marker is selected from the tables 1 to 20, 25 or 27, 29, 30, 32, 33, 35, 36, 38, 39, 41 or 42.

The assessment may be effected on the nucleic acid level such as by hybridization techniques or PCR or on the protein level such as by using antibody or aptamers based technologies.

Finally, the invention relates to a diagnostic composition comprising at least one nucleic acid molecule which is capable of specifically hybridizing to the mRNA corresponding to the marker gene of any of the appended tables. The nucleic acid

molecule may be an antisense DNA or RNA an RNAi molecule a siRNA molecule or the like inhibitory molecule capable of specifically blocking trnascription and/or translation and/or modification and/or localization of the RNA and/or protein corresponding to the marker gene.

The nucleic acid may also be a sense-strand nucleic acid e.g. RNA or preferably DNA which is capable of expressing the protein product of the marker gene, or a protein product of substantially similar activity, in a target cell into which it is introduced.

The invention further comprises pharmaceutical compositions comprising a compound capable of specifically binding to a protein or RNA corresponding to a marker of the invention as listed in any of the appended tables. The marker is preferably selected from the markers designated as particular preferred markers as described herein above . The compound is preferably a compound capable of inhibiting or increasing the function of the protein or of enhancing or decreasing translation of the RNA. The compound is preferably selected from aptameres, aptazynes, RNAzynes, antibodies, affybodies, trinextins, anticalins, or the like compounds. The effect of the compounds on the RNA may be tested by assaying for increased/decreased synthesis of the corresponding protein. The effect of the compounds on the protein may be assayed the testing the effect of the compound in an assay of the proteins function, which e.g. may be an anzymathic function. Alternatively, the effect may be tested by contacting a leukemic cell that expresses large amounts of such protein with the compound and assay cellular parameters associated with the leukemic state of the cell, such as cell growth, growth factor dependency and/or differentiation state of the cell.

In a further embodiment, the invention provides a method of determining wether a patient sample contains leukemia cells or other cells comprising the steps of

a) determining the expression profile of a group of markers in a patient sample and

b) concluding from the expression profile whether the patient sample contains leukemia cells or other cells, and optionally, to which subtype said leukemia cells belong, wherein

a subtype or a type of leukemia listed in table 28 b or c is identified, and a sensitivity and/or specificity of at least 80, 85, 88, 90, 92, 95, 97, 98, 99, 99.1, 99.2, 99.3, 99.4 or 99.5% is achieved, preferably using at least one marker of the group of markers listed in table 29 and/or 30.

In a further embodiment, the invention provides a method of determining wether a patient sample contains leukemia cells or other cells comprising the steps of

(a) determining the expression profile of a group of markers in a patient sample and

(b) concluding from the expression profile whether the patient sample contains leukemia cells or other cells, and optionally, to which subtype said leukemia cells belong, wherein

a subtype or a type of leukemia listed in table 31 b or c is identified, and a sensitivity and/or specificity of at least 80, 85, 88, 90, 92, 95, 97, 98, 99, 99.1, 99.2, 99.3, 99.4 or 99.5% is achieved, preferably using at least one marker of the group of markers listed in table 32 and/or 33.

In a further embodiment, the invention provides a method of determining wether a patient sample contains leukemia cells or other cells comprising the steps of

(a) determining the expression profile of a group of markers in a patient sample and

(b) concluding from the expression profile whether the patient sample contains leukemia cells or other cells, and optionally, to which subtype said leukemia cells belong, wherein

a subtype or a type of leukemia listed in table 34 b or c is identified, and a sensitivity and/or specificity of at least 80, 85, 88, 90, 92, 95, 97, 98, 99, 99.1, 99.2, 99.3, 99.4 or 99.5% is achieved, preferably using at least one marker of the group of markers listed in table 35 and/or 36.

In a further embodiment, the invention provides a method of determining wether a patient sample contains leukemia cells or other cells comprising the steps of

(a) -determining the expression profile of a group of markers in a patient sample and

(b) concluding from the expression profile whether the patient sample contains leukemia cells or other cells, and optionally, to which subtype said leukemia cells belong, wherein

a subtype or a type of leukemia listed in table 37 b or c is identified, and a sensitivity and/or specificity of at least 80, 85, 88, 90, 92, 95, 97, 98, 99, 99.1, 99.2, 99.3, 99.4 or 99.5% is achieved, preferably using at least one marker of the group of markers listed in table 38 and/or 39.

In a further embodiment, the invention provides a method of determining wether a patient sample contains leukemia cells or other cells comprising the steps of

(a) determining the expression profile of a group of markers in a patient sample and

(b) concluding from the expression profile whether the patient sample contains leukemia cells or other cells, and optionally, to which subtype said leukemia cells belong, wherein

a subtype or a type of leukemia listed in table 40 b or c is identified, and a sensitivity and/or specificity of at least 80, 85, 88, 90, 92, 95, 97, 98, 99, 99.1, 99.2, 99.3, 99.4 or 99.5% is achieved, preferably using at least one marker of the group of markers listed in table 41 and/or 42.

Description of the Figures

Figure 1a:	Principal Component Analysis
Figure 1b:	Hierarchical Cluster Analysis
Figure 2:	Classification Accuracy
Figures 3a, 3b1, 3b2:	PCA of AML data based on 312 genes Decision Trees according to I(E)
Figure 4:	Pair-wise Comparison of Normal BM and AML
Figure 5a:	Principal Component Analysis
Figure 5b:	Hierarchical Cluster Analysis
Figure 5c:	Pair-wise Comparison of Normal BM and ALL
Figure 6a:	Principal Component Analysis
Figure 6b:	Hierarchical Cluster Analysis
Figure 6c:	Pair-wise Comparison of Normal BM and CML
Figure 7a:	Principal Component Analysis
Figure 7b	Hierarchical Cluster Analysis
Figure 7c:	Pair-wise Comparison of Normal BM and CLL
Figure 8a:	Principal Component Analysis
Figure 8b:	Hierarchical Cluster Analysis
Figure 8c:	AML-WHO Classification
Figure 9a:	Principal Component Analysis
Figure 9b:	Hierarchical Cluster Analysis
Figure 9c:	Comparison of Normal BM versus Leukemia
Figure 10a:	Principal Component Analysis
Figure 10b:	Hierarchical Cluster Analysis
Figure 10c:	
Figure 11a	Accurate diagnosis of leukemia is accomplished in a two-step approach. First, samples are assigned to one of the major leukemia types or normal BM, respectively. Then, if positive for ALL or AML, further subclassification based on cytogenetically

	defined characteristics is proposed. In total 111 samples were analyzed by gene expression profiling and implemented in the development of different class prediction models: normal BM (n=8), CLL (n=8), CML (n=10), ALL (n=18), and AML (n=59). 18 ALL samples can further be characterized by B-lineage ALL samples positive for t(8;14) (n=3), t(9;22) (n=7), or t(11q23)/MLL (n=4) and T-lineage ALL (n=3), respectively. Additionally, one B-ALL sample showed an aberrant karyotype. 59 AML samples were comprized of normal karyotype (n=3), complex aberrant karyotype (n=4), trisomy 8 as sole abnormality (n=3), t(8;21) (n=9), t(15;17) (n=16), inv(16) (n=10), and t(11q23)/MLL (n=10). The latter four AML entities were additionally represented by each of the following t(8;21),+8 (n=1), t(15;17),+8 (n=2), and inv(16),+8 (n=1). Furthermore, some expression profiles were excluded for development of the classifier but subsequently tested for performance in diagnostical class assignments: normal BM (n=1), CLL (n=2), CML (n=2), ALL with t(4;11) (n=1), and AML with t(15;17) (n=2), respectively.
Figure 11b:	Hierarchical clustering of 55 AML samples (rows) versus 25 informative genes (columns). In total, 15 comparisons within the 5 groups were performed (pairwise and one-versus-all). Genes were selected for maximal accuracy and confidence based on a modified signal-to-noise (S2N) algorithm. The scaled gene expression levels are coded by intensity and shown on a scale from black (no expression) to bright red (highest expression). The AML subgroups 'other' (n=10), t(11q23)/MLL (n=10), inv(16) (n=10), t(8;21) (n=9), and t(15;17) (n=16) are colored according to their chromosomal aberrations. The minimal set of informative genes is given by HGNC approved symbols (not yet approved genes are marked by asterisks).
Figure 11c	Hierarchical clustering of 17 ALL samples (rows) versus 19 informative genes (columns). In total, 10 pairwise or OVA comparisons within the 4 groups were performed. Genes were selected for maximal accuracy and confidence based on a modified S2N algorithm. The scaled gene expression levels are coded by intensity and shown on a scale from black (no

	expression) to bright red (highest expression). The ALL subgroups t(11q23)/MLL (n=4), t(9;22) (n=7), t(8;14) (n=3), and T-ALL (n=3) are colored according to their characteristical chromosomal aberrations or immunophenotype. The minimal set of informative genes is given by HGNC approved symbols (asterisks mark not yet approved genes).
Figure 12a – 12i	Bar graphs of gene expression intensities for distinct leukemia types and subtypes. A short description indicates the respective classes which can be distinguished at each case.
Figure 13a.	Dot plot of expression levels for a particular gene in two groups (e.g. group1= normal samples, group2 = disease samples). Golub's decision limit to distinguish between group1 and group2, which is defined as the mean of μ_1 and μ_2 (μ_a: mean expression in group a), is not optimal, because the standard deviations of gene expression levels within the two groups are very different. In this case, a lower limit (e.g. maximum level in group1) would have been more appropriate to separate the two groups by means of gene expression levels.
Figure 13b	Accuracy and confidence for all-pairs and one-versus-all comparisons in a dataset consisting of 103 samples from 5 classes (A,B,C,D,E) using Golub's method and *diffgenes*. Both accuracy and confidence are higher with *diffgenes*.
Figure 14	Detailed characteristics of the 37 AML cases representing three defined cytogenetic aberrations corresponding to four cytomorphological subtypes according to FAB classification: inv(16)(p13q22)/AML M4eo, t(8;21)(q22;q22)/AML M2, and t(15;17)(q22;q12)/AML M3 or M3v. Diagnosis was proven by a) karyotype analysis, b) interphase-FISH (*CBFB, AML1* and *ETO, PML* and *RARA*), c) RT-PCR (*CBFB-MYH11, AML1-ETO, PML-RARA*), and d) cytomorphology.
Figure 15	Figure 15: Three cytogenetically defined AML subtypes with t(15;17), t(8;21) or inv(16) can be separated based on their gene expression profiles of 1,000 preselected genes. The three

	different subgroups form distinct clusters. For visualization in a two-dimensional plot the first two principal components were chosen as they captured most of the variation in the original data set. The subgroups are coloured according to their chromosomal aberrations, respectively
Figure 16	Hierarchical cluster analysis of the gene expression pattern of the set of 13 predictor genes as identified according to the adapted class prediction methodology introduced by Golub et al. The three distinct cytogenetic AML subgroups can clearly be separated based on their gene expression profiles. Each row represents a leukemia sample and each column a gene. The gene accession numbers are shown on the top. Varying expression levels are shown on a scale from black (no gene expression) to bright red (highest expression). The subgroups are coloured according to their chromosomal aberrations, respectively.
Figure 17	Schematic representation of the 15 decision trees (a to o) used in the multiple-tree classifier. Arrows indicate high (arrow up) or low (arrow down) expression, "0" and "+" denote absence or presence of a gene, respectively (e.g., in (a) the low expression of X96719 indicates AML with t(15;17) whereas the high expression of X96719 indicates AML with inv(16) or AML with t(8;21); the latter two entities are distinguished by X53742: lack of expression identifies AML with inv(16) and positive expression predicts AML with t(8;21)). The GenBank accession numbers are given for genes the expression level of which are used for decision. Nodes are represented as ovals and leaves as rectangles. Classes are referred to as t(15;17), t(8;21) or inv(16).
Figure 18	Based on a preselection of 82 genes morphologically different but cytogenetically identical AML subtypes M3 with t(15;17) and M3v with t(15;17) can be separated based on their gene expression profile. AML M3 samples are shown as green dots, AML M3v samples as blue dots, respectively.

Figure 19:	Correlations between protein expression levels and mRNA abundance. Expression levels were compared by Pearson´s correlation. Mean fluorescence intensity values obtained by flow cytometry were calculated for all events with fluorescence values higher than isotype controls using the CellQuest Pro software (Beckton Dickinson). Average fluorescence intensity values obtained by micorarray analyses were calculated by the Affymetrix software, Microarray Suite, Version 4.0.1.
Figure 20	Detailed characteristics of the 45 AML cases representing four defined recurrent cytogenetic abnormalities. Diagnosis was proven by a) karyo-type analysis, b) interphase-FISH, c) RT-PCR, and d) cytomorphology.
Fig. 21	Class separation by principal component analysis (PCA)
Fig. 22	Figure 3: PCA-Plot based on 39 informative genes. All leukemia samples could accurately be assigned to their corresponding cytogenetic subtype with 100% accuracies. To illustrate these results, a hierarchical clustering is shown (Fig. 4).
Fig. 23	Hierarchical clustering of 44 diagnostic AML samples and 8 normal BM samples (columns) versus 39 informative genes (rows). Gene expression levels are coded by intensity and represented on a scale from black (no expression) to bright red (highest expression).
Fig. 24 to 465	Bar graphs of gene expression intensities for distinct leukemia types and subtypes or normal bone marrow, respectively. Selected statistically significant genes are given by Affymetrix identification number and Human Gene Nomenclature Committee approved symbol (where available). A short description indicates the respective classes which can be distinguished at each case.

List of References

Breiman et al., Classification and regression try, Wadsworth & Brooks/Cole (Monterey)

Efron and Tibshirani, An introduction to the bootstrap (1993), Chapman & Hall (New York, London), pp. 237-247

Jolliffe, Principle Components Analysis (1986), Springer (New York)

Lockhart, D. J., et al., Nat Biotechnol 14 (1996) 1675-80

Sambrook, J., et al., in "Molecular Cloning: A Laboratory Manual" (1989), Eds. J. Sambrook, E. F. Fritsch and T. Maniatis, Cold Spring Harbour Laboratory Press, Cold Spring Harbour, NY

EP 0 237 362

The following examples, references, sequence listing and figures are provided to aid the understanding of the present invention, the true scope of which is set forth in the appended claims. It is understood that modifications can be made in the procedures set forth without departing from the spirit of the invention.

Examples

EXAMPLE 1

EXAMPLE 1 - General Methods

EXAMPLE 1 - (A) Selection and characterisation of Leukemia Samples

Bone marrow (BM) aspirates were taken at the time of the initial diagnostic biopsy and remaining material was immediately lysed in RLT buffer (Qiagen), frozen and stored at -80 C until preparation for gene expression analysis. For microarray analysis the GeneChip System (Affymetrix, Santa Clara, CA, USA) was used. The targets for GeneChip analysis were prepared according to the current Expression Analysis. Briefly, frozen lysates of the leukemia samples were thawed, homogenized (QIAshredder, Qiagen) and total RNA extracted (RNeasy Mini Kit, Qiagen).Normally 10 ug total RNA isolated from 1 x 107 cells was used as starting material in the subsequent cDNA-Synthesis using Oligo-dT-T7-Promotor Primer (cDNA synthesis Kit, Roche Molecular Biochemicals). The cDNA was purified by phenol-chlorophorm extraction and precipitated with 100% Ethanol over night. For detection of the hybridized target nucleic acid biotin-labeled ribonucleotides were incorporated during the in vitro transcription reaction (Enzo® BioArray™ HighYield™ RNA Transcript Labeling Kit, ENZO). After quantification of the purified cRNA (RNeasy Mini Kit, Qiagen), 15 ug were fragmented by alkaline treatment (200 mM Tris-acetate, pH 8.2, 500 mM potassium acetate, 150 mM magnesium acetate) and added to the hybridization cocktail sufficient for 5 hybridizations on standard GeneChip microarrays. Before expression profiling Test3 Probe Arrays (Affymetrix) were chosen for monitoring of the integrity of the

cRNA. Only labeled cRNA-cocktails which showed a ratio of the messured intensity of the 3' to the 5' end of the GAPDH gene less than 3.0 were selected for subsequent hybridization on HG-U95Av2 probe arrays (Affymetrix). Washing and staining the Probe arrays was performed as described (siehe Affymetrix-Original-Literatur (LOCKHART und LIPSHUTZ). The Affymetrix software (Microarray Suite, Version 4.0.1) extracted fluorescence intensities from each element on the arrays as detected by confocal laser scanning according to the manufacturers recommendations.

EXAMPLE 1 - (B) Data analysis

Class separation by principal component analysis and hierarchical cluster analysis: In a first step we reduced the dimensionality of the number of genes. Therefore we scaled the data from each array to a target intensity value 50 (Affymetrix Microarray Suite) in order to be able to perform inter-array comparisons. Then all data was analyzed using Significance Analysis of Microarrays (Multiclass Response, Stanford University) and we selected a distinct number of genes based on a permutations test. This reduced set of genes which showed to be significant then was analyzed using the public available Java application J-Express analysis tool (download at www.molmine.com). Principal Component Analysis and Hierarchical Cluster Analysis (parameters Cluster method: single linkage and Distance metric: euclidean) showed a clear separation of analyzed groups of samples e.g. healthy bone marrow versus leukemia.

EXAMPLE 1 - **(C) Identification of differentially expressed genes according to Golub et al. (Science 1999 Oct 15;286(5439):531-7)**

A previously described (Science 1999 Oct 15;286(5439):531-7) was modified to reduce the number of candidate genes that could distinguish between our leukemic samples of interest. In a first step the raw data was scaled using Affymetrix software (target intensity 50 for all genes). To avoid division by zero or

negative numbers as occuring due to the current expression algorithm (Affymetrix) we set all average intensities of 20 or less to 20. Briefly, for a more detailed gene expression profiling we applied the data analysis method according to Golub et al. using weighted voting. In a first step gene expression levels were log-transformed with a cut-off value set at 20 units. To assess the significance of selected genes we performed a leave-one-out cross-validation. Only those genes were considered important which were contained in all cross validation classificators. To determine the association between genes by chance we performed a permutation test (100 cycles). Because the number of informative genes, which are able to discriminate between samples, is unknown, we applied the Golub method for different numbers of informative genes (range: 10-200). The minimal set of genes which provided optimal classification accuracy was selected to avoid overfitting.

EXAMPLE 2

EXAMPLE 2 - Identification of genes, the aberrant expression of which is associated with a particular leukemia subtype

Monitoring the gene expression level of thousands of mRNA transcripts simultaneously in one experiment is the key technology to find out the specific genes which allow the subsequent development of a class prediction model. We therefore used the Affymetrix oligonucleotide microarray technology (GeneChip® Instrument System) to obtain gene expression profiles of each individual clinical sample of interest. The HG-U95Av2 probe arrays gave us information about the relative mRNA abundance of about 12,000 full length human genes which are represented on these high-density oligonucleotide microarrays.

In total, 8 bone marrow samples of healthy volunteers and leukemia patients were investigated. Five different types of bioinformatic calculations were performed.

EXAMPLE 2 **(I) Three distinct genetic subtypes of AML**

Three defined cytogenetic aberrations t(8;21)(q22;q22) (n=9), t(15;17)(q22;q12) (n=16) and M4eo with inv(16) (p13q22) (n=10) corresponding to the 4 FAB-subtypes AML M2, M3 or M3v and M4eo, respectively. After we obtained bone marrow aspirates from 35 untreated patients with newly diagnosed AML, all cases were characterized by cytomorphology, cytogenetics and by molecular genetics (Fig. 1). AML subtypes M3 and M3v both carry the same chromosomal aberration but differ in morphological aspects like nuclear configuration, granulation and clinical aspects white blood cell count (WBC), respectively. In all cases, these balanced abnormalities were confirmed by fluorescence in-situ hybridization. The corresponding fusion transcript was detected by RT-PCR and/or quantitative real time PCR. The median age of the patients was 53 years (range, 19-82 years) and did not differ between the respective groups. The median WBC count was 17.0 G/l (range, 0.8-168.0 G/l) and was strikingly lower in patients with AML M3 as compared to all other patients.

EXAMPLE 2 - Methods used

EXAMPLE 2 - (A) Selection and characterisation of Leukemia Samples

We obtained bone marrow (BM) aspirates from 37 AML patients standing for four morphological and three underlying cytogenetic subgroups that were sent to the Laboratory of Leukemia Diagnostics (LFL) for central diagnosis within the German AMLCG study (Klinikum Grosshadern, Munich, Germany). They were selected for this study on the basis of several criteria. It was mandatory that none of the patients had been treated. All samples, exclusively newly diagnosed in our laboratory, had to be well characterized as de novo AML and diagnosis had been proven by cytomorphology, cytogenetics, flow cytometry and molecular genetics in every single case. All samples for gene expression analysis were taken at the time of the initial diagnostic biopsy when remaining material was immediately lysed in

RLT buffer (Qiagen), frozen and stored at -80 C until preparation for gene expression analysis.

EXAMPLE 2 - (B) Microarray experiments

For microarray analysis the GeneChip System (Affymetrix, Santa Clara, CA, USA) was used. The targets for GeneChip analysis were prepared according to the current Expression Analysis Technical Manual. Briefly, frozen lysates of the leukemia samples were thawed, homogenized (QIAshredder, Qiagen) and total RNA extracted (RNeasy Mini Kit, Qiagen). Normally 10 ug total RNA isolated from 1 x 107 cells was used as starting material in the subsequent cDNA-Synthesis using Oligo-dT-T7-Promotor Primer (cDNA synthesis Kit, Roche Molecular Biochemicals). The cDNA was purified by phenol-chlorophorm extraction and precipitated with 100% Ethanol over night. For detection of the hybridized target nucleic acid biotin-labeled ribonucleotides were incorporated during the in vitro transcription reaction (Enzo® BioArray™ HighYield™ RNA Transcript Labeling Kit, ENZO). After quantification of the purified cRNA (RNeasy Mini Kit, Qiagen), 15 ug were fragmented by alkaline treatment (200 mM Tris-acetate, pH 8.2, 500 mM potassium acetate, 150 mM magnesium acetate) and added to the hybridization cocktail sufficient for 5 hybridizations on standard GeneChip microarrays. Before expression profiling Test3 Probe Arrays (Affymetrix) were chosen for monitoring of the integrity of the cRNA. Only labeled cRNA-cocktails which showed a ratio of the measured intensity of the 3' to the 5' end of the GAPDH gene less than 3 were selected for hybridization on HG-U95Av2 probe arrays (Affymetrix). Washing and staining the Probe arrays was performed as described. The Affymetrix software (Microarray Suite, Version 4.0.1) extracted fluorescence intensities from each element on the arrays as detected by confocal laser scanning according to the manufacturers recommendations.

EXAMPLE 2 - (C) Class separation by principal component analysis and hierarchical cluster analysis

In a first step we reduced the dimensionality of the number of genes. Therefore we scaled the data from ,each array to a target intensity value 50 (Affymetrix Microarray Suite) in order to be able to perform inter-array comparisons. Then all data was analyzed using Significance Analysis of Microarrays (Multiclass Response, Stanford University) and we selected 580 genes based on a permutations test. This reduced set of genes which showed to be significant then was analyzed using the public available Java application J-Express analysis tool (download at www.molmine.com). Principal Component Analysis and Hierarchical Cluster Analysis (parameters Cluster method: single linkage and Distance metric: euclidean) showed a clear separation of analyzed groups of samples e.g. healthy bone marrow versus leukemia.

EXAMPLE 2 - (D) Identification of differentially expressed genes according to Golub

This analysis was carried cut as described in Example 1 (C) above. Briefly, classification of tumor samples was achieved by using a set of samples whose class had been already determined. This set was called training set. By using the oligonucleotide microarrays (Lockhart, D. J., et al., Nat Biotechnol 14 (1996) 1675-80), the, transcript levels in training set samples were measured for those genes that were represented on the microarray. The values for "transcripticn strength" were determined by averaging the values of a set of probes which were compared to a set of nearly identical probes containing a single mismatch. This was performed by using; methods provided by the oligonucleotide array of Affymetrix Inc.

EXAMPLE 2 - (E) Principle Components Analysis, Classifier and DecisionsTrees

In order to obtain comparable values between different samples, they had to be standardized first. The method followed that described (Lockhart, D. J., et al., Nat Biotechnol 14 (1996) 1675-80), except that correcting for (additive) background

had been omitted. In brief, the data from one of the samples were declared to serve as a "standard", and the values from all other samples were adapted to this standard. For every possible comparison to this standard, a set of "reliable" values was determined by calculating the correlation coefficient for a series of intervals of increasing length. The lower bound of reliability was the bound of the interval that had a correlation coefficient less than or equal to the smaller intervals. From all reliable values, 2 (logarithmized) correction factor was calculated by computing the median of the differences of the logarithmic values. Values that were zero or negative prior to taking the logarithm were not taken into account.

The obtained data matrix contained values from one sample per column. The gene expression profile across all samples for one gene or gene fragment represented on the oligonucleotide microarray was contained in a row of the matrix. To allow for rapid calculation of the classifier and to reduce memory usage, certain genes were pre-selected from the set of all genes represented on the array. The following criteria were applied:

Formula (1):

$$\sum_{i=1}^{k} |\mu_i - \overline{\mu}| - \sum_{i=1}^{k} \sigma_i > 0$$

Formula (2):

$$t < \frac{\sum_{i=1}^{k} |\mu_i - \overline{\mu}|}{\sum_{i=1}^{k} \sigma_i}$$

μ_1 refers to the average of the i-th class (i=l,...,k), μ to the total average, σ_i to the standard deviation of the i-th class and t to an arbitrary treshold ≤ 1. Selection by these methods resulted typically in a reduction in the number of genes by a factor of 10-30. To check the quality of the selection procedure, the first two principal components (Jolliffe, Principle Components Analysis (1986), Springer (New York)) for the samples were plotted. This allowed to judge whether or not a rigorous discrimination was possible between the different classes.

For construction of the classifier, decision trees (Breiman et al., Classification and regression try, Wadsworth & Brooks/Cole (Monterey)) were used. Simple decision trees that discriminate between n classes by using only transcription levels for (n-l) genes were used. They were trained and the selected genes were the discarded from the original data set. A new tree was constructed by using the truncated data set and the entire procedure was iterated until a predetermined number of trees was reached. The optimal number of trees could be estimated by counting the number of misclassifications of classifiers built from different numbers of trees. For this, an independent data set of cross-validation had to be used. The final vote of the multi-classifier was obtained by applying a vote-by-majority rule to the predictions o f the contained trees. In the example of the present invention 15 decision trees had been used for the multi-classifier. This allowed perfect classification of 100% of the samples, discriminating between classes that were given by chromosomal aberrations. To estimate generalization properties, i.e. how accurate the classifier may perform on samples that have not been used for training, cross-validation had been used (Efron and Tibshirani, An introduction to the bootstrap (1993), Chapman & Hall (New York, London), pp. 237-247).

EXAMPLE 2 - Results (Golub Method)

From this point of view it was found that a set of 17 genes was sufficient to distinguish distinct AML subtypes from each other with high precision (Tables 1). The classification model was able to identify the 4 morphologically and 3 cytogenetically and molecular biological different subtypes AML with t(8;21), with t(15;17), and with inv(l6) (Figures la-b, 2).

In conclusion by comparison of gene expression profiles of AML samples (3 tested genetic subtypes t(8;21), t(15;17) and inv(16)) genes could be identified which allowed a differentiation between each individual AML subtype in detail could be shown for the first time that these distinct abnormalities on the genomic level relate to a specific gene expression pattern. In other words, in the experimental setting the knowledge of the expression status of' these designated genes was sufficient to predict the genetic abnormality and allows the diagnosis of specific genetically defined subtypes of AML (Table 1).

Results of methods described in I(E) are shown in Table 2 and Figures 3a + b, 1/2 and 4.

EXAMPLE 2 - **II) Pair-wise comparisons between normal bone marrow, AML, ALL, CML, and CLL**: By pair-wise comparisons gene expression profiles of 8 cases of normal bone marrow, 48 AML, 9 ALL, 8 CML, and 7 CLL were evaluated. These led to the identification of subtype-specific genes (Tables 3-12. Figs. 5a-c, 6a-c, 7a-c, 8a-c).

EXAMPLE 2 - **III) AML classified according to WHO proposal**

To allow classification of AML subtypes according to the new WHO proposal we used the gene expression profiles of four genetically defined AML subtypes (t(8;21) n= 9; t(15;17) n= 18; inv(16) n= 10; 11q23/MLL aberrations n= 11). This led to the identification of subtype-specific genes (Table 13, Figs. 9a-c).

EXAMPLE 2 - **IV) Normal bone marrow *versus* distinct genetic subtypes of AML**: We used the gene expression profiles of normal bone marrow (n=8) and of four genetically defined AML subtypes (t(8;21) n= 9; t(15;17) n= 18; inv(16) n= 10; 11q23/MLL aberrations n= 10). This led to the identification of genes that allow the distinction between normal bone marrow and each of the four AML subtypes (Table 14).

EXAMPLE 2 - **V) Identification of genes specifically separating normal bone marrow, AML, ALL, CML, and CLL:** : We used the gene expression profiles of normal bone marrow (n=8) and of AML (n=48), ALL (n = 9), CML (n = 8), and CLL (n =7). This led to the identification of xx genes that allow the distinction between normal bone marrow and each of the four leukemia subtypes (Table 15, Figures 10a-c).

Example 3: Gene expression profiling provides a global and robust diagnostic tool for leukemia

Example 3- Introduction

The expression profiles of 12,600 genes were analyzed in 103 patients suffering from chronic myeloid leukemia (CML), chronic lymphoid leukemia (CLL), acute lymphoblastic leukemia (ALL), and acute myeloid leukemia (AML). A set of 71 genes shown in table 16 to 19 was identified as the minimal set necessary to accurately diagnose prognostically relevant leukemia subtypes and to distinguish these from normal bone marrow (BM, n=8). Thus, microarray technology is a suitable method for diagnosis of leukemia.

Today, the classification of hematological malignancies according to the WHO criteria describes chronic myeloid leukemia (CML), chronic lymphoid (CLL), acute lymphoblastic (ALL), and acute myeloid leukemia (AML). Within the latter two several prognostically relevant subtypes are established (see example 4). This subclassification is based on genetic abnormalities of the leukemic blasts associated with different prognoses and becomes increasingly important to guide therapy. Thus, the development of new, specific treatment approaches requires the precise identification of these subtypes that may benefit from individual therapeutic protocols. It has already been shown that the development of drugs targeting molecular aberrations can dramatically improve outcome. The introduction of all-trans retinoic acid (ATRA) into the treatment of AML with t(15;17)(q22;q11-12) has improved outcome from about 50% to 80% long-term survivors (1). In CML patients imatinib, a designed molecule that inhibits the t(9;22)(q34;q11) specific chimeric tyrosine kinase BCR-ABL, induces dramatically higher response rates as compared to conventional drugs (2). To fully take advantage of specific treatment options a precise identification of distinct leukemia subtypes is mandatory. However, standard diagnostics of leukemia using a combination of complementary methods is expensive, time-consuming, and requires experienced specialists.

Since its introduction, microarrays have been promising tools for basic research. With regard to leukemia, the pivotal discrimination of unselected ALL and AML samples based on their gene expression signatures inspired numerous studies (3). Recently, subtypes of childhood ALL could be correlated to specific gene

expression profiles leading to both marker genes suitable for initial diagnostics and canditates as predictors for outcome (Yeoh, Eng-Juh. pediadric ALL expression profiling *Cancer Cell*, 2002). Additionally, novel entities in hematological malignancies could be identified based on their distinct expression pattern as has been shown for multiple myeloma, large cell lymphoma, and childhood ALL (4-6). In example 4, it is demonstrated that cytogenetically defined AML subtypes can be correlated to specific gene expression profiles (see example 4). AML FAB M2 with t(8;21)(q22;q22), FAB M3/M3v with t(15;17)(q22;q11-12), or M4eo with inv(16)(p13q22) could be classified based on a minimal set of 13 genes. These genes belong to a large variety of different functional classes including members of signaling pathways, cell surface antigens, as well as plasma glycoproteins. Several genes are known to be involved in cytoskeletal structure, transcriptional processes, or have not yet further been functionally described.

Here, gene expression profiles of 103 leukemia patients were acquired representing 11 groups and eight normal BM donors to designate leukemia-specific genes which build up the basis for a novel diagnostic tool. Following the aims of Golub, who introduced the cancer class prediction methodology (3, 7), all four major leukemia types were analyzed and also included cytogenetically defined subgroups of AML and ALL as described in the WHO classification of leukemia, respectively (Fig. 11a). All patient samples were thoroughly characterized combining cytomorphology, cytogenetics, immunophenotyping, and molecular genetics. This was a prerequisite to obtain disease-specific gene expression profiles for each entity. We used Affymetrix expression probe arrays HG-U95Av2 to interrogate the mRNA abundance of approximately 12,600 transcripts. In order to identify genes suitable for a leukemia prediction classifier we applied a slightly modified prediction methodology as introduced by Golub [see (Note1_Golub method)]. A minimal set of candidate genes had to show both maximal classification accuracy and maximal confidence. Accuracy of the classifiers was determined by permutation-based neighborhood analysis [see (Note2_ leave-one-out crossvalidation)]. Additional information about the absolute differences of expression intensities and further descriptions of all candidate genes can be found in the supporting online material.

In a first step, based on 23 informative genes the samples were assigned to either normal BM, CLL, CML, ALL, or AML, respectively (Table 22; Description of Table 22: Classification scheme for 4 major leukemia types and normal BM. Matrices delineate distribution of actual leukemia types as compared with predicted types from pairwise comparisons. Class assignment can be based on the expression profiles of 23 genes. Except for pairwise comparison of AML versus ALL, all cases can be predicted accurately in leave-one-out cross validation with 100% accuracy and strong confidence. For each pairwise comparison the minimal set of informative genes is represented by approved HUGO Gene Nomenclature Committee (HGNC) symbols. Not yet approved genes are marked by asterisks.). In 9/10 pairwise comparisons all samples were classified correctly (335 individual assignments; 100% accuracy). In one comparison (AML versus ALL) 75/77 samples were classified correctly resulting in an accuracy of 97%. Two ALL samples were misclassified as AML. This may be due to the heterogenity of both groups (n=18 versus n=59) causing noise in the expression data.

For each pairwise comparison a set of discriminative genes is disclosed in table 20 whereby the gene names can be found in table 21. The most discriminative and informative genes are marked by asterisks in table 20 and are the 71 genes shown in table 16 to 19

In detail, we found phospholipidscramblase 1 (*PLSCR1*) to be lower expressed in AML and ALL as compared to normal BM. *PLSCR1* encodes for a plasma membrane protein and has been proposed to play a role in transbilayer migration of phospholipids and in recognition and phagocytic clearance of injured, aged, or apoptotic cells (8). The biologic effects of interferon-alpha may be mediated by *PLSCR1* which is markedly upregulated by interferon (9, 10). We also observed that *LEF-1* was absent in myeloid leukemias but highly expressed in lymphoid leukemias. *LEF-1* was shown to be mitogenic and important for cell survival in pro-B cells (11). The B-cell specific coactivator of octamer binding transcription factors, *POU2AF1*, plays an important role in the antigen-driven stages of B cell activation and maturation and discriminates between AML and CLL (12). *MSF* has been reported to be a translocation partner of the mixed-lineage leukemia gene (MLL) in AML and was able to separate AML from ALL (13). Likewise, *OS-9*, not yet further

functionally described except for amplification in osteosarcomas, was differentially expressed between AML and ALL (14). *HLA-DMB* plays a critical role in antigen presentation by catalyzing the release of class II HLA-associated invariant chain binding sites for acquisition of antigenic peptides (15). It is known that lymphocytes in CLL express high levels of class II antigens whereas mature myeloid leukemias are e.g. HLA-DR negative (16, 17). Therefore, the differential expression of *HLA-DMB* in CML as compared to CLL illustrates well the differential expression of cell surface MHC class II molecules. *NCOA1* plays a critical role in STAT3 and STAT6 pathways and was expressed higher in CLL as compared to ALL suggesting an inhibitory effect of STAT6-mediated transactivation in CLL (18). A member of the tumor necrosis factor receptor family, whose surface expression has already been reported in CLL (19), *CD27*, was identified to assign samples either ALL or CLL. We also detected *LCN2* that was shown to be a modulator of inflammation regulated by interleukin-9 with highest expression in CML samples (20). *IRF4*, an immune system-restricted interferon regulatory factor that is required for lymphocyte activation showed no expression in CML while it was expressed in normal BM. Recently, an increase of IRF4 levels in CML patients demonstrated an association with a good response to interferon-alpha therapy (21). Several other proteins (*DEFA3*, *SGP28*, *CAMP*, *CLC*) are known to be stored in the granules of neutrophils and allowed assignment of leukemic samples to the CML type if highly expressed (22-25).

The second step of our approach was to build up a classifier for the identification of AML subtypes genetically defined according to the WHO classification, i.e. AML with t(8;21), with t(15;17) with inv(16), and with 11q23-translocations involving the MLL gene, respectively. In addition, a category 'other' was analyzed comprizing AML with normal karyotype (n=3), AML with complex aberrant karyotype (n=4), and AML with trisomy 8 as sole abnormality (n=3), respectively. A set of 25 most informative genes was identified based on pairwise comparisons and one-versus-all (OVA) comparisons. None of these genes had already been identified for the classification of the four leukemia types and normal BM as described above. As shown in Figure 11b, distinct AML subgroups cluster together due to homogeneous expression profiles. This classification model showed 100% classification accuracies in 14/15 comparisons (440 individual assignments). In

one OVA comparison, 'other' versus all other AML, 54/55 samples were assigned correctly. The missclassification of one sample may also reflect the large heterogenity of both groups.

The following genes were identified in OVA comparisons and discriminate distinct AML subtypes. The gene most valuable for prediction of AML M4eo with inv(16) was *MYH11*. Its higher expression as compared to all other AML most probably is due to hybridization of the M4eo-specific fusion transcripts *CBFβ-MYH11* to corresponding *MYH11*-oligonucleotides represented on the microarray (26). Likewise, the higher expression of *CBFA2T1* (formerly *ETO*) in AML with t(8;21) may be due to a similar effect of hybridization of the subtype-specific *AML1-ETO* fusion transcript (27). Another highly characteristic gene for t(8;21) positive AML was *POU4F1*, which has been described to play an important role in retinal ganglion cell differentiation and has recently been shown to confer an oncogenic potential when co-transfected with *H-RAS* (28). Furthermore, it was shown to be highly expressed in neuro-epithelioma and ewing sarcomas (29). Another member of this transcription factor family, *POU2F2*, was able to discriminate between t(11q23)/MLL versus group 'other'. A related gene, *POU2AF1*, has recently been reported to be underexpressed in acute leukemia with t(11q23)/MLL-rearrangement (5). The most informative genes in our approach discriminating AML with t(11q23)/MLL-rearrangement from all other AML subtypes were *SOCS-2* and *MBNL*. Generally, *SOCS-2* shows a higher expression level in AML with t(11q23)/MLL-rearrangement and is known to play a role in cytokine-induced signaling pathways (30). Similarly, *MBNL* shows a higher expression in AML with t(11q23)/MLL-rearrangement as compared to all other AML samples. Its encoded protein as well as other MBL family members are localized in the nucleus and share a Cys3His zinc finger motif (31). MBL proteins occur in several isoforms due to alternative splicing (32) and may have different functions as has been shown for *HOX* genes (33). *HOXA9* has been reported to be highly expressed in leukemia with MLL-rearrangements (5). In contrast, expression of *HOXB5* is characteristic of AML group 'other' as compared to all other AML subtypes in our data. The most important genes discriminating AML with t(15;17) from all other AML subtypes were *ARGHGAP4* and *CTSW*. ARGHGAP4 is predominantly expressed in hematopoietic cells but showed a lower expression level in AML with t(15;17) as

compared to all other AML subtypes. It encodes a member of signaling proteins involved in regulation of small GTP-binding proteins of the RAS-superfamily, which themselves play an important role in cell cycle and apoptosis (34). *CTSW* encodes for a recently described papain-like cysteine protease, which is predominantly expressed in NK cells and to a lesser extent in cytotoxic lymphocytes. It may represent a putative component of the endoplasmatic reticulum resident proteolytic machinery (35). A survey about the expression levels of genes in the AML subtypes can be found in Fig. 12a-d

Subclassification of ALL comprizing the three B-lineage groups ALL with t(9;22), with t(4;11), or with t(8;14) was analyzed next and compared with T-lineage ALL expression profiles. All samples were classified correctly on the basis of 19 genes (Fig. 11c). This set included *TRB*, which was already described to distinguish between CLL and CML (Table 22).

In detail, the genes encoding for the T cell receptor beta subunit and T cell surface CD3 delta chain (*TRB*, *CD3D*) were identified as highly indicative of T-ALL as compared to both ALL with t(9;22) and all other ALL subtypes. This is in line with standard diagnostics of T-ALL by immunophenotyping where these antigens comprize the most specific ones (36). Similarly, *MME* (formerly *CD10*) was highly expressed in ALL with t(9;22) only. This on the one hand may reflect that t(9;22) is observed in common-ALL and in pre-B ALL only. On the other hand, this data again demonstrates that the gene used for diagnostic purposes in flow cytometry, *MME*, may be highly indicative of these ALL subtypes in comparisons to the more immature B-lineage ALL, i.e. pro-B ALL, as well as the mature B-ALL and the T-ALL. Furthermore, the identification of connective tissue growth factor (*CTGF*) as a specific marker for ALL with t(4;11) adds to previous data demonstrating its increased gene expression in childhood ALL in general (37). The glucocorticoid receptor beta has been shown to be highly expressed in ALL with t(4;11) but not in t(9;22) positive ALL. This is in line with the particularly poor prognosis of the latter subtype since response to corticoid therapy is one of the most powerful prognostic factors in ALL (38, 39). In addition, we speculate that new treatment options may be realized for T-ALL based on the high expression of adenosine deaminase (*ADA*) in this subtype. Inhibitors of *ADA* have been shown to be effective in indolent T-cell lymphomas but have not yet been evaluated in T-ALL (40). One

cytokine differentially expressed between t(8;14) positive ALL and T-lineage ALL was *SCYA3*. We recommend testing the monitoring of its protein expression as a supplemental antigen useful for immunophenotypical identification of t(8;14) positive ALL. Finally, in ALL carrying t(4;11) *v-myb* is highly expressed and may thus be involved in the pathogenesis of this subtype. In general, a role of v-myb has been described for the transformation of myelomonocytic cells (41). A survey about the expression levels of genes in the AML subtypes can be found in Fig. 12e-12i.

At least, we intended to separate t(9;22) positive from t(9;22) negative ALL. Our data contained two genes encoding for *ADCY3* and the hypothetical protein *KIAA1013* which were sufficient for the 100% correct assignments of 18 analyzed cases. Both genes showed a higher expression in t(9;22) positive as compared to t(9;22) negative ALL. Additionally, distinguishing B-lineage from T-lineage ALL, *CD3D* and *TRB* repeatedly showed their usefulness as T-ALL marker genes as already described in Figure 11c (18/18 correct individual assignments).

Generally, chromosomal aberrations are strongly associated with morphological characteristics. However, there are two chromosomal aberrations which are observed in both myeloid and lymphatic neoplasms, i.e. t(11q23)/MLL and the t(9;22). The t(9;22) occurs in ALL and CML, and t(11q23)/MLL is observed in ALL and AML, respectively. Analyzing gene expression signatures of both t(9;22) positive ALL and CML we identified two genes, which allowed 17/17 correct lineage assignments. *CD74* plays a critical role in MHC class II antigen processing and demonstrated a lower expression in t(9;22) positive CML (42). This may also explain the relationship between the low MHC class II antigen presentation in CML in general and fits well to the recognized lower *HLA-DMB* expression in CML as compared to CLL (Table 1). *CAPN3* is a member of the papain superfamily and was higher expressed in CML discriminating them from t(9;22) positive ALL [see (Note_ 38894_g_at)].

In addition, our results indicate that the expression signatures of two genes, *CD24* and *CTGF*, are sufficient for 14/14 correct assignments of the t(11q23)/MLL positive leukemias either to ALL or to AML. Thus, in both scenarios lineage assignment can be accomplished based on specific gene expression signatures despite the same underlying chromosomal aberrations.

Taken together, these data demonstrate the utility of gene expression profiling using microarrays for diagnosis of leukemia. In total, 11 different leukemia entities could clearly be distinguished from each other and from normal BM, respectively. These leukemias are associated with highly differing prognoses and require specific treatment strategies. By performing these analyses on a single platform requiring basic molecular biological methods, this approach guarantees broad access to high-quality diagnosis of leukemia. The robust gene expression analysis with high diagnostic accuracy can substitute the combination of cytomorphology, cytogenetics, immunophenotyping, and molecular biological methods used today. Compared to the combination of methods used so far, this approach also reduces costs. In order to introduce diagnostical genomics into routine clinical practice, prospective trials in parallel to conventional methods are necessary to prove the reliability in a large cohort of patients. Furthermore, gene expression patterns will allow the additional subclassification of leukemia especially in subtypes with no specific cytogenetic markers and the identification of deregulated master genes within distinct leukemia entities can guide the way to new therapeutic approaches.

Reference List of Example 3

1. Fenaux, P., Le Deley, M. C., Castaigne, S., Archimbaud, E., Chomienne, C., Link, H., Guerci, A., Duarte, M., Daniel, M. T., Bowen, D. *et al.* (1993) *Blood* **82**, 3241-3249.
2. Kantarjian, H., Sawyers, C., Hochhaus, A., Guilhot, F., Schiffer, C., Gambacorti-Passerini, C., Niederwieser, D., Resta, D., Capdeville, R., Zoellner, U. *et al.* (2002) *N. Engl. J. Med.* **346**, 645-652.
3. Golub, T. R., Slonim, D. K., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J. P., Coller, H., Loh, M. L., Downing, J. R., Caligiuri, M. A. *et al.* (1999) *Science* **286**, 531-537.
4. Alizadeh, A. A., Eisen, M. B., Davis, R. E., Ma, C., Lossos, I. S., Rosenwald, A., Boldrick, J. C., Sabet, H., Tran, T., Yu, X. *et al.* (2000) *Nature* **403**, 503-511.
5. Armstrong, S. A., Staunton, J. E., Silverman, L. B., Pieters, R., den Boer, M. L., Minden, M. D., Sallan, S. E., Lander, E. S., Golub, T. R. & Korsmeyer, S. J. (2002) *Nat. Genet.* **30**, 41-47.

6. Zhan, F., Hardin, J., Kordsmeier, B., Bumm, K., Zheng, M., Tian, E., Sanderson, R., Yang, Y., Wilson, C., Zangari, M. *et al.* (2002) *Blood* **99**, 1745-1757.

7. Pomeroy, S. L., Tamayo, P., Gaasenbeek, M., Sturla, L. M., Angelo, M., McLaughlin, M. E., Kim, J. Y., Goumnerova, L. C., Black, P. M., Lau, C. *et al.* (2002) *Nature* **415**, 436-442.

8. Fadok, V. A., Bratton, D. L., Rose, D. M., Pearson, A., Ezekewitz, R. A. & Henson, P. M. (2000) *Nature* **405**, 85-90.

9. Der, S. D., Zhou, A., Williams, B. R. & Silverman, R. H. (1998) *Proc. Natl. Acad. Sci. U. S. A* **95**, 15623-15628.

10. Zhou, Q., Zhao, J., Al Zoghaibi, F., Zhou, A., Wiedmer, T., Silverman, R. H. & Sims, P. J. (2000) *Blood* **95**, 2593-2599.

11. Reya, T., O'Riordan, M., Okamura, R., Devaney, E., Willert, K., Nusse, R. & Grosschedl, R. (2000) *Immunity.* **13**, 15-24.

12. Nielsen, P. J., Georgiev, O., Lorenz, B. & Schaffner, W. (1996) *Eur. J. Immunol.* **26**, 3214-3218.

13. Osaka, M., Rowley, J. D. & Zeleznik, L. (1999) *Proc. Natl. Acad. Sci. U. S. A* **96**, 6428-6433.

14. Su, Y. A., Hutter, C. M., Trent, J. M. & Meltzer, P. S. (1996) *Mol. Carcinog.* **15**, 270-275.

15. Van Kaer, L. (2001) *Immunol. Res.* **23**, 205-214.

16. Marti, G. E., Zenger, V., Caproaso, N. E., Brown, M., Washington, G. C., Carter, P., Schechter, G. & Noguchi, P. (1989) *Anal. Quant. Cytol. Histol.* **11**, 315-323.

17. Newman, R. A. & Greaves, M. F. (1982) *Clin. Exp. Immunol.* **50**, 41-50.

18. Litterst, C. M. & Pfitzner, E. (2001) *J. Biol. Chem.* **276**, 45713-45721.

19. van Oers, M. H., Pals, S. T., Evers, L. M., van der Schoot, C. E., Koopman, G., Bonfrer, J. M., Hintzen, R. Q., dem Borne, A. E. & van Lier, R. A. (1993) *Blood* **82**, 3430-3436.

20. Cowland, J. B. & Borregaard, N. (1997) *Genomics* **45**, 17-23.

21. Schmidt, M., Hochhaus, A., Konig-Merediz, S. A., Brendel, C., Proba, J., Hoppe, G. J., Wittig, B., Ehninger, G., Hehlmann, R. & Neubauer, A. (2000) *J. Clin. Oncol.* **18**, 3331-3338.

22. Oren, A. & Taylor, J. M. (1995) *J. Lab Clin. Med.* **125**, 340-347.

23. Cowland, J. B. & Borregaard, N. (1999) *J. Leukoc. Biol.* **66**, 989-995.

24. Dvorak, A. M., Letourneau, L., Login, G. R., Weller, P. F. & Ackerman, S. J. (1988) *Blood* **72**, 150-158.

25. Gudmundsson, G. H., Agerberth, B., Odeberg, J., Bergman, T., Olsson, B. & Salcedo, R. (1996) *Eur. J. Biochem.* **238**, 325-332.

26. van der Reijden, B. A., Dauwerse, J. G., Wessels, J. W., Beverstock, G. C., Hagemeijer, A., van Ommen, G. J. & Breuning, M. H. (1993) *Blood* **82**, 2948-2952.

27. Downing, J. R., Head, D. R., Curcio-Brint, A. M., Hulshof, M. G., Motroni, T. A., Raimondi, S. C., Carroll, A. J., Drabkin, H. A., Willman, C., Theil, K. S. *et al.* (1993) *Blood* **81**, 2860-2865.

28. Liu, W., Khare, S. L., Liang, X., Peters, M. A., Liu, X., Cepko, C. L. & Xiang, M. (2000) *Development* **127**, 3237-3247.

29. Leblond-Francillard, M., Picon, A., Bertagna, X. & de Keyzer, Y. (1997) *J. Clin. Endocrinol. Metab* **82**, 89-94.

30. Krebs, D. L. & Hilton, D. J. (2001) *Stem Cells* **19**, 378-387.

31. Begemann, G., Paricio, N., Artero, R., Kiss, I., Perez-Alonso, M. & Mlodzik, M. (1997) *Development* **124**, 4321-4331.

32. Fardaei, M., Rogers, M. T., Thorpe, H. M., Larkin, K., Hamshere, M. G., Harper, P. S. & Brook, J. D. (2002) *Hum. Mol. Genet.* **11**, 805-814.

33. van Oostveen, J., Bijl, J., Raaphorst, F., Walboomers, J. & Meijer, C. (1999) *Leukemia* **13**, 1675-1690.

34. Tribioli, C., Droetto, S., Bione, S., Cesareni, G., Torrisi, M. R., Lotti, L. V., Lanfrancone, L., Toniolo, D. & Pelicci, P. (1996) *Proc. Natl. Acad. Sci. U. S. A* **93**, 695-699.

35. Wex, T., Buhling, F., Wex, H., Gunther, D., Malfertheiner, P., Weber, E. & Bromme, D. (2001) *J. Immunol.* **167**, 2172-2178.

36. Campana, D. & Behm, F. G. (2000) *J. Immunol. Methods* **243**, 59-75.

37. Vorwerk, P., Wex, H., Hohmann, B., Mohnike, K., Schmidt, U. & Mittler, U. (2002) *Mol. Pathol.* **55**, 40-45.

38. Gleissner, B., Gokbuget, N., Bartram, C. R., Janssen, B., Rieder, H., Janssen, J. W., Fonatsch, C., Heyll, A., Voliotis, D., Beck, J. *et al.* (2002) *Blood* **99**, 1536-1543.

39. Panzer-Grumayer, E. R., Schneider, M., Panzer, S., Fasching, K. & Gadner, H. (2000) *Blood* **95**, 790-794.

40. O'Brien, S., Kurzrock, R., Duvic, M., Kantarjian, H., Stass, S., Robertson, L. E., Estey, E., Pierce, S. & Keating, M. J. (1994) *Blood* **84**, 733-738.

41. Dvorakova, M., Kralova, J., Karafiat, V., Bartunek, P. & Dvorak, M. (2001) *Blood Cells Mol. Dis.* **27**, 437-445.

42. Villadangos, J. A. (2001) *Mol. Immunol.* **38**, 329-346.

Notes of Example 3

[see (Note1_Golub method)]

When comparing two groups of microarray experiments, Golub's method sorts the genes with respect to the signal-to-noise ratio of gene x: $S_x = (\mu_1-\mu_2)/(\sigma_1+\sigma_2)$, where μ_k and σ_k denote the mean expression and standard deviation of gene x in group k. According to a specified number of "informative" genes (e.g. 20) the best discriminating genes are selected. For each informative gene a decision limit is calculated as $b_x = (\mu_1+\mu_2)/2$. To classify a new sample of an independent test set, the gene expression levels of informative genes are taken and for each gene x and sample y a so-called vote is calculated as $V_x = S_x\,(g_x^y - b_x)$, where g_x^y denotes expression level of gene x in sample y. The votes of all informative genes are summed up ("weighted voting") and depending upon the sign of this sum the new sample is classified as group 1 or group 2. The *confidence* in the prediction is calculated as $|\Sigma V_x / \Sigma |V_x||$.

However, the decision limit proposed by Golub does not provide optimal classification accuracy in all situations. Importantly, when the standard deviation of expression levels within the two groups are very different, the decision limit is biased towards the group with the higher standard deviation. A decision limit for a particular gene can be considered optimal, if it achieves maximum classification accuracy for a given dataset. By determining systematically classification accuracies for a set of possible decision limits, an optimal decision limit can be

calculated. We selected an optimal decision limit from the following set of decision limits L_x: $L_x = \{ (g_x^y + g_x^{y-1})/2 \mid 1 < y <= n \}$ where g_x^y denotes expression level of gene x in sample y, n denotes the total number of samples in the training set.

Additionally, we applied an heuristic approach to select a minimal set of discriminative genes, which provides maximum classification accuracy in leave-one-out-crossvalidation. We applied for a given set of 20 informative genes weighted voting as described above and the classification accuracy was calculated by crossvalidation. Therefore, our algorithm consists of the following steps: (i) Calculate the top 20 discriminating genes according to the signal-to-noise ratio. (ii) Calculate classification accuracy and confidence based on optimal decision limits for each of the top 20 genes (iii) Select the gene which provides best classification accuracy and confidence out of step 2. (iv) Test for each of the remaining 19 genes, whether adding this gene to the model improves accuracy and confidence; if the gene improves accuracy and confidence, it is added to the weighted voting model, otherwise it is discarded.

In detail, this method can be described as follows:

Example 3 - Subheading to Note1_Golub method: Abstracts

Differentially expressed genes can potentially be used in medical diagnostics, if the gene expression patterns are reliable and specific for a particular disease. *diffgenes* is a program to identify differentially expressed genes in microarray experiments. Its algorithm is based on the method proposed by Golub, but contains two improvements: an optimized decision limit per gene and a minimal set of discriminative genes.

The new method was applied to a human dataset from the domain of cancer research consisting of 103 microarrays with 12625 genes each. *diffgenes* outperforms Golub's method clearly both in terms of accuracy and confidence of classifications. The biological validation of the results is facilitated, because *diffgenes* identifies a very small number of candidate genes (typically < 5). Microarray datasets can be analyzed with *diffgenes* on the Internet at http://martin-dugas.de/diffgenes/

Example 3 - Subheading to Note1_Golub method: Introduction

Microarrays are used in ongoing research to characterize disease processes on a molecular level. Gene expression analysis enables to identify new subtypes within known diseases with prognostic relevance for the patients [Alizadeh 2000].

For interpretation of the wealth of data - more than 10.000 parameters per experiment - it is advisable to integrate microarray data with detailed clinical information. For applications in medical diagnostics, significant associations between gene expression profiles and sample groups resulting in classification accuracies in the range of 70 - 80% are not sufficient; for diagnostic purposes at least 95% classification accuracy is required.

If a certain disease is characterized by a specific gene product, e.g. a pathologic fusion gene, a precise measurement of the expression of this particular gene should be a reliable marker for the disease. Therefore in a diagnostic setting, very few and specific genes would be desirable.

However, for many diseases the precise molecular pathogenesis is not yet known. In addition, the function of many genes on currently available microarrays like Affymetrix GeneChip[R] is still unclear.

Therefore microarray data should be analyzed and interpreted carefully. By integration of data from different diagnostic modalities (morphology, PCR, FISH, clinical data) the biological plausibility and consistency of microarray data can be verified.

Example 3 - **Subheading to Note1 Golub method: Methods**

Example 3 - **Subheading to Note1 Golub method: Golub's method**

When comparing two groups of microarray experiments, Golub's method sorts the genes with respect to the signal-to-noise ratio of gene x: $S_x = (\mu_1-\mu_2)/(\sigma_1+\sigma_2)$, where μ_k and σ_k denote the mean expression and standard deviation of gene x in group k.

According to a specified number of "informative" genes (e.g. 20) the best discriminating genes are selected. For each informative gene a decision limit is calculated as $b_x = (\mu_1+\mu_2)/2$. To classify a new sample of an independent test set,

the gene expression levels of informative genes are taken and for each gene x and sample y a so-called vote is calculated as $V_x = S_x (g_x^y - b_x)$, where g_x^y denotes expression level of gene x in sample y. The votes of all informative genes are summed up ("weighted voting") and depending upon the sign of this sum the new sample is classified as group 1 or group 2. The *confidence* in the prediction is calculated as $|\Sigma V_x / \Sigma |V_x| |$. To assess the significance of each gene, a permutation test is performed, which determines signal-to-noise ratios when class labels are permuted randomly. To assess the robustness of the classifier, a leave-one-out crossvalidation is performed. *Accuracy* is the rate of correctly classified test samples. Further details are contained in [Golub 1999], [Pomeroy 2002, Supplement].

Example 3 - Subheading to Note1 Golub method: An optimized decision limit

The decision limit proposed by Golub does not provide optimal classification accuracy in all situations. As can be seen in Figure 13a, when the standard deviation of expression levels within the two groups are very different, the decision limit is biased towards the group with the higher standard deviation.

A decision limit for a particular gene can be considered optimal, if it achieves maximum classification accuracy for a given dataset. By determining systematically classification accuracies for a set of possible decision limits, an optimal decision limit can be calculated. The *diffgenes* program selects an optimal decision limit from the following set of decision limits L_x:

$$L_x = \{ (g_x^y + g_x^{y-1})/2 \mid 1 < y <= n \}$$

where g_x^y denotes expression level of gene x in sample y, n denotes the total number of samples in the training set.

Example 3 - Subheading to Note1_Golub method: A minimal set of discriminative genes

Golubs method selects an arbitrary number of "informative" genes to discriminate between two classes of samples according to their signal-to-noise ratio, typically in the range of 10 to 50 genes. Choosing too many genes carries the risk of overfitting, which causes poor generalization features of the model. Therefore *diffgenes* applies an heuristic approach to select a minimal set of discriminative genes, which provides maximum classification accuracy in leave-one-out-crossvalidation. I.e. for a given set of genes weighted voting as described by Golub is applied and the classification accuracy is calculated by crossvalidation.

The *diffgenes* algorithm consists of the following steps:

1. Calculate the top 20 discriminating genes according to the signal-to-noise ratio

2. Calculate classification accuracy and confidence based on optimal decision limits for each of the top 20 genes

3. Select the gene which provides best classification accuracy and confidence out of step 2.

4. Test for each of the remaining 19 genes, whether adding this gene to the model improves accuracy and confidence; if the gene improves accuracy and confidence, it is added to the weighted voting model, otherwise it is discarded.

Example 3 - <u>Subheading to Note1 Golub method: Results</u>

The method was applied to a new human dataset from the domain of cancer research consisting of 103 Affymetrix Genechip(R) microarrays with 12625 genes each. Table 23 presents an analysis of 18 samples class A versus 85 samples class non-A (Description of Table 23: Analysis of 18 samples class A versus 85 samples class non-A. On the left the analysis according to Golub is presented for 20 informative genes. The crossvalidation accuracy is 0,87, confidence 0,77. Samples, where crossvalidation failed, are listed. For each gene signal to noise ratio, p-value (significance obtained from permutation test) and decision limit are provided. On the right the same data set is analyzed using *diffgenes*. By selection

of 3 genes (marked with asterisks) out of the top 20 genes and selecting optimized decision limits, the crossvalidation accuracy reaches 0,96, confidence 0,88.). Based on 20 informative genes Golub's method results in a crossvalidation accuracy of 0;87 (confidence 0,77); *diffgenes* achieves with three genes out of the top 20 set a crossvalidation accuracy of 0,96 (confidence 0,88). The same analysis was performed for one versus all (OVA) and all pairs (AP) comparisons in this dataset consisting of 5 different classes. Figure 13b presents accuracy and confidence obtained by both methods: *diffgenes* outperforms Golub's method clearly both in terms of accuracy and confidence of classifications. The same comparative approach was applied to two datasets in cardiology and cell biology consisting of 44 and 67 microarrays. The results concerning Golub's method and *diffgenes* were very similar (data not shown).

Example 3 - Subheading to Note1 Golub method: **Discussion**

There are two major challenges in the analysis of microarray data: the number of variables (genes) is much higher than the number of individual samples and the correlation structure of the parameters is widely unknown.

Golub's method to analyse microarray data has been applied to important medical datasets [Armstrong 2002]. Recently many different approaches have been applied to microarray data: Classical statistical techniques like ANOVA with adjustment for multiple testing, significance analysis of microarrays (SAM) [Tusher 2001] , selection of discriminative genes with support vector machines (SVM), neural networks and many more. This indicates that the underlying problem is important and non-trivial; a comparison of different methods is needed. Robustness of the generated mathematical models is an important issue, therefore bootstrap procedures and permutation tests are applied.

For medical diagnostics differentially expressed genes are of interest, but the sensitivity and specificity for particular diseases must be validated prospectively in larger patient cohorts. *diffgenes* is an extension of Golub's method to improve classification accuracy, which is very relevant in a diagnostic setting. The optimized decision limit plays an important role, because the situation presented in Figure 13a is quite common in biological contexts: group 1 represents samples, where the expression of gene x is repressed while gene x is activated in group 2. The biological validation of the results is facilitated, because *diffgenes* identifies a very small number of candidate genes (typically < 5).

Emphasis must be placed on verification of results by other diagnostic procedures, because the selected "important" genes are not only dependent on the statistics procedure, but also on the preprocessing of data. In our setting by integration of microarray analysis with other laboratory modalities (morphology, cytogenetics, molecular genetics, immunphenotyping) and clinical data the plausibility and consistency of results could be evaluated, therefore we are optimistic, that the demanding requirements for medical diagnostics can be fulfilled with microarray technology in the near future.

Example 3 - Subheading to Note1 Golub method: References

Alizadeh AA, Eisen MB, Davis RE, et al. (2000) Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature 403(6769):503-11

Armstrong SA, Staunton JE, Silverman LB, et al. (2002) MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia. Nature Genetics 1:41-7

Golub TR, Slonim DK, Tamayo P, et al. (1999) Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science 286(5439):531-7

Pomeroy SL, Tamayo P, Gaasenbeek M, et al. (2002) Prediction of central nervous system embryonal tumour outcome based on gene expression. Nature 415(6870):436-42

Tusher VG, Tibshirani R, Chu G (2001) Significance analysis of microarrays applied to the ionizing radiation response. PNAS 98: 5116-5121

EXAMPLE 3 - [see (Note2_ leave-one-out crossvalidation)]

To assess the significance of each gene, a permutation test is performed, which determines signal-to-noise ratios when class labels are permuted randomly. To assess the robustness of the classifier, a leave-one-out crossvalidation is performed. *Accuracy* is the rate of correctly classified test samples.

EXAMPLE 3 - [see (Note_ 38894_g_at)]

The second top-ranked gene was represented by the Affymetrix probe set identifier: 38894_g_a. However, no clear gene assignment was possible for this informative prove set. Therefore, CAPN3 was chosen.

Example 4: PNAS

EXAMPLE 4 - ABSTRACT

Acute myeloid leukemia (AML) is a heterogeneous group of genetically defined diseases. Their classification is important with regard to prognosis and treatment. We performed microarray analyses for gene expression profiling on bone marrow samples of 37 patients with newly diagnosed AML. All cases had either of the distinct subtypes AML M2 with t(8;21), AML M3 or M3v with t(15;17), or AML M4eo with inv(16). Diagnosis was established by cytomorphology, cytogenetics, fluorescence-in-situ hybridization, and RT-PCR in every sample. By using two different strategies for microarray data analyses, this study for the first time revealed a unique correlation between AML-specific cytogenetic aberrations and gene expression profiles.

EXAMPLE 4 - INTRODUCTION

Acute myeloid leukemia (AML) is a heterogeneous group of diseases with respect to biology and clinical course. Since the introduction of the FAB-classification in 1976 diagnosis and classification have been based on cytomorphology and cytochemistry(1). As other techniques like immunophenotyping, cytogenetics, and molecular genetics contributed to the definition of AML subtypes the FAB-classification was updated. In 1999 the WHO classification for tumors of hematopoietic and lymphoid tissues was proposed. In an attempt to define biologically homogeneous entities which have clinical relevance morphologic, immunophenotypic, genetic and clinical features were incorporated(2, 3).

For optimal treatment approaches both a precise diagnosis and prognostic parameters that determine response to therapy and survival are needed. So far, the karyotype of the AML blasts is the most important independent prognostic factor. A favorable outcome under currently used treatment regimens with cure rates from 50% up to 85% was observed in several studies in patients with a) t(8;21)(q22;q22) occuring mostly in FAB subtype AML M2, b) inv(16)(p13q22) associated with AML M4eo and c) t(15;17)(q22;q11-12) associated with AML M3 and AML M3v(4-6). In contrast, chromosome aberrations with an unfavorable clinical course are -5/del(5q), -7/del(7q), inv(3)/t(3;3) and complex aberrant

karyotypes with cure rates of less than 10%(7, 8). The remainder AML patients are assigned to a prognostically intermediate group. This latter group is very heterogeneous because it includes patients with a normal karyotype as well as those with rare chromosome aberrations and yet unknown prognostic impact.

Besides their prognostic impact genetic aberrations are involved in the pathogenesis of leukemia. While for unbalanced cytogenetic aberrations the heterogeneous pathogenetic mechanisms have not yet conclusively been determined, several studies provide strong evidence for the central pathogenetic role of leukemia-specific fusion genes that are generated by the above mentioned balanced abnormalities(9-12). Therefore it can be postulated that AML with balanced abnormalities most probably display a homogeneous gene expression profile and thus are promising candidates for microarray analyses.

In a pivotal study, gene expression profiles were analyzed in bone marrow samples of 27 ALL and 11 AML. A set of 50 genes out of 6,817 analyzed genes was sufficient to discriminate ALL and AML. By leave-one-out cross-validation it was possible to correctly classify 36 out of 38 acute leukemia cases. A class predictor could automatically determine new leukemia cases out of an independent test set as belonging to the myeloid or the lymphoid lineage. Thus, these results demonstrated the possibility of cancer classification based on gene expression profiling(13). In a further approach comparing AML with trisomy 8 and AML with normal karyotype expression profiling revealed fundamental biological differences in AML with isolated trisomy 8 and normal cytogenetics(14). More recently, acute lymphoblastic leukemias (ALL) with translocations involving the *MLL* gene could be separated from ALL cases without MLL translocations and from cases with AML by gene expression profiling(15).

The aim of our investigation was to answer the question whether a leukemia specific genotype is associated with a distinct gene expression profile. Therefore, we analyzed three distinct genetic subtypes of acute myeloid leukemia: t(8;21)(q22;q22), inv(16)(p13q22) and t(15;17)(q22;q12) which lead to subtype specific fusion genes *AML1-ETO*, *CBFB-MYH11* and *PML-RARA*, respectively. They are specifically associated with four distinct morphological subtypes according to the FAB-classification: AML M2, AML M4eo, AML M3 and AML M3v(16-18). We performed microarray analyses on a cohort of leukemia samples (*n*=37) and applied several methodologies to evaluate genes which allowed an assignment to the corresponding type of cytogenetic aberration for classification.

This is the first time that AML-specific cytogenetic aberrations can be correlated with corresponding gene expression profiles and vice versa.

EXAMPLE 4- METHODS

Example 4- Selection and characterization of leukemia samples

For this investigation we selected bone marrow (BM) samples from 37 AML patients representing four morphological and three underlying cytogenetic subgroups. All cases were sent for reference diagnostics to our laboratory and registered in our leukemia database(19). Samples were received either locally or by overnight mail. All samples were newly diagnosed *de novo* AML and were characterized by cytomorphology, cytogenetics, FISH, and molecular genetics in each case. Gene expression analyses were performed on cells remaining from the diagnostic samples. Samples had been lysed immediately, frozen and were stored at -80°C from one to 34 months until preparation for gene expression analysis.

Example 4- Cytomorphology

Analysis was based on May-Grünwald-Giemsa stain, myeloperoxidase reaction, and non-specific esterase reaction using alpha-naphthyl-acetate. All staining from bone marrow and blood was performed routinely according to standard procedures(20). The cytomorphologic diagnosis followed the criteria of the FAB classification and the new WHO classification(1, 3, 18).

Example 4- Cytogenetics

Chromosome analyses were performed on bone marrow or peripheral blood samples according to standard protocols(21). Metaphases were analyzed for G-bands using a modified GAG-banding technique as described elsewhere(22). Twenty to 25 metaphase cells were analyzed. The chromosomes were interpreted according to the International System for Human Cytogenetic Nomenclature(23).

Example 4- Fluorescence in situ hybridization (FISH) on interphase nuclei

FISH was performed on interphase nuclei on bone marrow smears or on slides prepared for cytogenetic analysis. For interphase-FISH at least 100 interphase nuclei were evaluated. FISH was carried out using commercially available *AML1-ETO*, *PML-RARA* and *CBFB* probes (VYSIS, Downers Grove, Il, USA). The signals were evaluated with an Axioskop^R (Zeiss, Jena, Germany). For documentation the analyzing system ISIS^R (MetaSystems, Altlussheim, Germany) was used.

Example 4- RNA isolation and Reverse-transcription-polymerase-chain-reaction (RT-PCR)

Mononuclear cells were isolated by a Ficoll gradient separation. 1x10^7 cells were lysed in RLT-buffer (Qiagen, Hilden, Germany) and total RNA was extracted with a RNeasy-kit (Qiagen) according to the manufacturers instructions. RNA was eluted in 50 μl of elution buffer.

Five μl of the total RNA, an equivalent quantity of 1x10^6 cells or about 1 μg of RNA were reversely transcribed in a 40 μl reaction using 300 U of Superscript^R (LifeTechnologies, Karlsruhe, Germany) and random hexamers (Pharmacia, Freiburg, Germany).

PCR for the specific *AML1-ETO*, *CBFB-MYH11*, or *PML-RARA* fusion transcripts were performed as has been described(24). For each sample an *ABL* specific RT-PCR was performed to control the integrity of RNA using primers ABL5′: 5′-GGCCAGTAGCATCTGACTTTG-3′ and ABL3′: 5′-ATGGTACCAGGAGTGTTTCTCC-3′. Strict precautions were taken to prevent contamination. Water instead of cDNA was included as a blank sample in each experiment. Amplification products were analyzed on 1.5% agarose gels stained with ethidium bromide.

Example 4- Microarray experiments

For microarray analysis the GeneChip® System (Affymetrix, Santa Clara, California) was used. The targets for GeneChip® analysis were prepared

according to the current Expression Analysis Technical Manual. Briefly, lysates of the leukemia samples were homogenized (QIAshredder, Qiagen, Hilden, Germany) and total RNA extracted (RNeasy Mini Kit, Qiagen). Normally, 10 μg total RNA isolated from 1x10^7 cells was used as starting material in the subsequent cDNA-synthesis using oligo[$(dT)_{24}$T7promotor]$_{65}$ primer (cDNA Synthesis System, Roche Diagnostics, Mannheim, Germany). The cDNA was purified by phenol:chlorophorm:IAA extraction (Ambion, Austin, Texas) and acetate/ethanol precipitated over night. For detection of the hybridized target nucleic acid biotin-labeled ribonucleotides were incorporated during the *in vitro* transcription (Enzo® BioArray™ HighYield™ RNA Transcript Labeling Kit, ENZO, Farmingdale, USA). After quantification of the purified cRNA (RNeasy Mini Kit, Qiagen), 15 μg was fragmented by alkaline treatment (200 mM Tris-acetate, pH 8.2, 500 mM potassium acetate, 150 mM magnesium acetate) and added to the hybridization cocktail sufficient for 5 hybridizations on standard GeneChip® microarrays. Before hybridization onto U95Av2, Test3 microarrays (Affymetrix) were chosen for monitoring of the integrity of the cRNA. Washing and staining of the probe arrays were performed according to the current protocols (Micro_1v1, EukGE-WS2v2). The Affymetrix software (Microarray Suite, Version 4.0.1) extracted fluorescence intensities from each element on the microarrays as detected by confocal laser scanning according to the manufacturers recommendations. Thirty-two out of 37 hybridization cocktails demonstrated high quality cRNA characteristics (Test3 probe arrays: 3'/5' ratio of *GAPDH* probe sets ≤3.0) and were selected for building up class prediction models.

Example 4- Class separation by principal component analysis

Potential clusters corresponding to the genetic subgroups were visualized applying a two-step approach. The data were scaled from each array to a target intensity value 50 (Affymetrix Microarray Suite 4.0.1) in order to be able to perform inter-array comparisons. All data were permutated 100 cycles using the multiclass response parameter of the Significance Analysis of Microarrays algorithm (SAM)(25) (http://www-stat.stanford.edu/~tibs/SAM/index.html). The total set of 12,600 genes was reduced to the significant differentially expressed genes. In a second step, the reduced set of genes was prepared for principal component analysis (PCA) and analyzed with J-Express(26) (http://www.molmine.com/). For

visualization in a two-dimensional plot we chose the first two principal components as they captured most of the variation in the original data set.

Example 4- Class prediction by weighted voting(13)

We adapted a previously described method to reduce the number of candidate genes that could distinguish between the three different cytogenetic AML subgroups(13). Briefly, to avoid division by zero or negative numbers as occurs due to the expression algorithm (Affymetrix Microarray Suite 4.0.1) we set all average fluorescence intensities of 1 or less to 1. Then, gene expression levels were log-transformed. Performing pairwise comparisons (A vs. B), for each gene g P(g,c) values and votes (defined by: P(g,c)=(m1(g)-m2(g))/(s1(g)+s2(g))) were calculated based on mean expression levels (m) and standard deviations (s) in the respective cytogenetic subgroup. Subsequently, votes were summed and prediction strength (PS) values reflected the margin of victory in the direction of either cytogenetic group A or B of the pairwise comparison. PS values range between 0 and 1, values >0.45 demonstrate significance (according to the permutation test). The relevance of selected genes was assessed by performing leave-one-out cross-validation. Only those genes that were contained in all cross validation classifiers were considered important. To determine a random association between genes we performed a permutation test (100 cycles). Because the number of informative genes, which are required to discriminate between samples, is unknown, we applied this method for different numbers of informative genes (range: 2 to 200). The minimal set of genes which provided optimal classification accuracy together with the highest prediction strength was selected to avoid overfitting. To visualize the identified genes and check their suitability for class separation a hierarchical cluster analysis was performed utilizing J-Express(26) (cluster method: average linkage; distance metric: euclidean). The accuracy of this class prediction model was validated on an independent test set of five cases of AML not fulfilling the cRNA high quality criterion as outlined above.

Example 4- Multiple-tree classifier

As basic units in this classifier, classification trees are used(27-29). The optimal number of trees has been determined to be 15 (data not shown). Class votes of these trees are aggregated by a vote-by-majority rule. The classifier was fed with gene expression intensity values from a set of 973 genes that had been chosen based on their r statistic:

$$r=\frac{\sum_{i=1}^{k}|\mu_i-\overline{\mu}|}{\sum_{i=1}^{k}\sigma_i}$$

where μ_i refers to the class averages, $\overline{\mu}$ to the overall average, σ_i to the within-class standard deviation, and summation is carried out over all k classes. The threshold was set to $r>0.75$. Classification trees were constructed as follows: tree building was performed while restricting trees to contain no more than n-1 nodes to discriminate between n classes. The C5.0 algorithm was used(28). The variables (gene expression intensities) used for tree construction were eliminated from the data set, and a new tree was calculated based on the truncated data set. This procedure was iterated until the predetermined number of trees had been reached. The accuracy of the multiple-tree classifier was estimated by 10-fold cross validation(30) and on an independent test set of data from 5 bone marrow aspirates, where the quality of the corresponding cRNA preparation was slightly lower than the high quality standards required for the training set.

EXAMPLE 4 - RESULTS

Example 4- Characterization of leukemia samples

We investigated 37 AML cases representing three defined cytogenetic aberrations corresponding to four FAB subtypes: t(8;21)(q22;q22)/AML M2 (n=9), t(15;17)(q22;q12)/AML M3 or AML M3v (n=10, n=8), and inv(16)(p13q22)/AML M4eo (n=10). All cases were characterized by cytomorphology, cytogenetics, FISH, and RT-PCR (Fig. 14). All cases with AML and t(8;21) had AML M2, all with AML and inv(16) had AML M4eo, ten cases with AML and t(15;17) had AML M3, and eight cases with AML and t(15;17) had AML M3v. All patients showed these balanced abnormalities as the sole karyotype change. Using FISH analysis, more than 65% of cells demonstrated the specific signal constellation. The respective fusion transcripts were detected by RT-PCR in all samples. The median age of all

patients was 53 years (range, 19-82 years; male:female=15:22) and did not differ between the respective groups. AML subtypes M3 and M3v both carry the same chromosomal aberration but differ in morphological aspects like nuclear configuration, granulation, and clinical aspects like white blood cell count (WBC). The median WBC count was 20,000/μl (range, 800-168,000/μl) and was strikingly lower in patients with AML M3 as compared to all other patients (median, 6,200 vs. 36,500/μl, P=0.0002).

Example 4- Microarray analyses

The gene expression profiles of 37 AML samples were evaluated. Thirty-two hybridization cocktails demonstrated high quality cRNA characteristics (Test3 probe arrays: 3'/5' ratio of GAPDH probe sets ≤3.0) and were selected for building class prediction models: t(8;21)/AML M2 (n=7), t(15;17)/AML M3 or M3v (n=9, n=7), and inv(16)/AML M4eo (n=9). Five cases were primarily excluded (3'/5' ratios ranging between 3.9 and 5.4, see Methods) and were used for subsequent validations of the class prediction models: t(8;21)/AML M2 (n=2), t(15;17)/AML M3 or M3v (n=1, n=1), and inv(16)/AML M4eo (n=1).

Example 4- Class separation by principal component analysis

In order to visualize clusters corresponding to the three underlying genetic subgroups we applied a two-step approach. Based on a permutation test (100 permutations) we correlated our expression data to the three different cytogenetic parameters(25). We obtained 1000 significant genes. By principal component analysis we were able to clearly separate the three distinct chromosomal aberrations t(8;21), t(15;17), and inv(16) (Fig. 15)(26). These data suggest that genetically defined AML subtypes can be specified and identified based on their gene expression profiles.

Example 4- Class prediction by weighted voting(13)

In order to identify the genes which enable the accurate discrimination of these subgroups, we applied the data analysis methodology introduced by Golub et al.(13). We selected the minimal set of genes which provided optimal classification accuracy together with the highest prediction strength to avoid overfitting. Thirteen genes were sufficient to separate these AML subtypes with high precision (Table 24; Table 24 shows that a minimal set of 13 genes (GenBank accession numbers are given) is sufficient for accurate class prediction with optimal classification accuracy and highest prediction strength. Comparisons (A vs. B) were performed either between two distinct subtypes or between one distinct subtype and all other subtypes (=remainder), respectively. As calculated from pairwise comparisons, positive P(g,c) values indicate a higher expression in first class listed, negative P(g,c) values a higher expression in second class listed, respectively). GenBank accession numbers and detailed descriptions of the genes are given in table 25 (Table 25: Thirty-six genes separate accurately three distinct cytogenetic AML subtypes. GenBank accession numbers, approved human gene nomenclature symbol (*=not approved) and description of the function are presented. Six genes are included in the minimal set of both weighted voting according to Golub et al.(13) (total=13) and multiple-tree classifiers (total=29).

All 32 clinical samples could be assigned to their corresponding cytogenetic subtype with best accuracy in leave-one-out cross-validation (1.0). Prediction strength values ranged from 0.91 to 0.98 (Table 24). To illustrate these results we applied hierarchical clustering(31). The resulting dendrogram clearly demonstrates the capacity of this subset of genes to separate all AML cases according to their cytogenetic aberration (Fig. 16). This demonstrates that class prediction of a chromosomal aberration in AML is feasable solely based on gene expression data.

For external validation, we tested whether primarily excluded samples could also be accurately assigned to their specific cytogenetic category. Despite their non-optimal cRNA quality, all 5 cases were correctly classified with high prediction strength (0.76,1.00,1.00,1.00,1.00).

Example 4- Class prediction by multiple-tree models

As a second and independent methodological approach we developed a multiple-tree classifier to separate the three genetically defined subtypes based on the expression level of a minimal set of genes. In short, we computed classification trees to discriminate between the different AML subclasses. To avoid overfitting of a singular tree model, we computed a multiple-tree model using an iteratively reduced set of genes. For each tree, we used only those genes that have not been used by the previously computed classification tree. The procedure is stopped when a predetermined number of trees has been reached. For this study, the optimal number of trees was calculated to be 15. The votes of the 15 trees were aggregated by a vote-by-majority rule. Equal votes for two of the three classes were counted as misclassification.

The classifier utilized the expression values of 29 genes (*MYH11* was identified twice by two different probe sets; Table 25) to discriminate between three classes, namely samples displaying t(15;17), t(8;21), and inv(16) (Fig. 17). The accuracy on the training set (*n*=32) was 100%, and on the independent test set (*n*=5) 100%. The average accuracy in ten-fold cross validation was 94%.

In summary, we identified 36 genes using two independent methodologies for class prediction in AML (Table 25). Six genes were described in both calculations, seven were found exclusively in the minimal set according to Golub et al.(13), and another 23 genes using multiple-tree classifiers.

Example 4- Correlation of phenotype and gene expression profile

We were able to demonstrate striking correlations between genotype and gene expression profiles in three genetically defined subgroups of AML. In addition, we answered the question, whether the cytogenetically identical AML with t(15;17) but appearing with two different phenotypes, AML M3 or AML M3v (Fig. 14), can also be separated by different gene expression patterns. We used 100-fold permutation of M3 (*n*=10) and M3v (*n*=8) data followed by principal component analysis and hierarchical cluster analysis based on 82 informative genes (data not shown). Separation into the corresponding two morphologically defined FAB subtypes M3 and M3v was possible in all cases (Fig. 18) and suggests also a close correlation between phenotype and gene expression profile.

EXAMPLE 4 - DISCUSSION

This is the first study to demonstrate an unequivocal association between disease-specific genetic alterations and distinct gene expression profiles. For each of the three analyzed clearly defined subtypes of AML (t(8;21), t(15;17), inv(16)) patterns of gene expression were identified that were homogeneous within all samples of the respective subgroups but clearly differed between these three subgroups. The analyzed samples represent disease subtypes that are specifically defined on the genetic and the phenotypic level by conventional diagnostics including cytomorphology, cytogenetics, and molecular genetics.

By applying two independent approaches for the analysis of microarray data, the present study demonstrates that AML samples from previously defined subtypes(3) can be classified adequately on the basis of gene expression profiles. It is intriguing that there is both sufficient coherence in gene expression within and difference between these subtypes to classify them with high accuracy even though the samples derive from the same myeloid cell lineage.

In order to correlate gene expression with cytogenetics Virtaneva et al. compared the expression status of 6,606 genes of AML blasts with normal cytogenetics and trisomy 8 as the sole abnormality. While in this study normal CD34+ cells clustered into a distinct group, AML with trisomy 8 and AML with normal karyotype intercalated with each other. Microarray analyses showed an overall increased gene expression of genes located on chromosome 8 suggesting a gene-dosage effect(14). AML with trisomy 8 is heterogeneous on the phenotypic level as it occurs in different FAB subtypes. In contrast, AML with t(15;17), inv(16) and t(8;21) show a very close correlation to distinct morphological subtypes. Furthermore, trisomy 8 is probably not a primary, disease-defining aberration leading to AML as it also occurs in addition to a variety of different cytogenetic and molecular genetic abnormalities(32, 33). In contrast to this study, Armstrong et al. compared samples of the more homogeneous group of ALL with *MLL* translocations to ALL without *MLL* translocations and to AML(15). They demonstrated that ALL with *MLL* translocations comprizes a distinct disease which can be classified robustly by gene expression profiling.

The main focus of the present analyses was the assessment of the differences between three highly characterized subgroups of AML defined by specific primary

chromosome aberrations. As anticipated, it was shown that AML with t(8;21) and AML with inv(16), which both involve alterations of the core binding factor-complex, are more related to each other as compared to AML with t(15;17)(34). Both phenotypically different subtypes of AML with t(15;17), AML M3 and AML M3v, cluster within one area. In an additional analysis, also the latter two subtypes were separated from each other based on their gene expression profiles. This data suggests the existence of further genetic and not yet identified alterations leading to the different phenotypes of AML M3 and AML M3v. One possible candidate gene is *FLT3* which is mutated more frequently in AML M3v than in AML M3 (67% vs. 19%, *P*=0.001)(35).

Several studies confirmed that gene expression profiles can be used for class prediction. This has been shown for acute leukemias, round blue cell tumors, and malignant melanomas(13, 36-38) as well as for different types of solid tumors using multi-class cancer classification(39). While the selection of different subgroups in these studies was performed using exclusively phenotypic criteria, other studies were based on genetically defined entities(40, 41). In the present study not only the discrimination of the three genetically defined AML subgroups was accomplished but also all these cases of AML were separated from normal bone marrow (data not shown)(42).

To develop a classifier two independent approaches were applied. While classification by weighted voting according to Golub et al.(13) allows the discrimination between the three classes based on a minimal set of 13 genes, the multiple-tree classifier utilizes 30 genes. As indicated by cross-validation, generalization properties are excellent for the multiple-tree classifier, i.e. it is likely to perform equally well on new, unseen samples. Furthermore, it can be easily extended to more than the three subclasses described in the present study.

Our classifiers contained genes already known to be primarily involved in the pathogenesis of the respective entities, namely *MYH11*(43) and *ETO*(44). Presumably, the detection of overexpression of *MYH11* in inv(16) cases and of *ETO* in t(8;21) cases relates to the detection of the fusion gene transcripts rather than of the wild type transcripts. The other genes identified belong to various functional categories. Their potential pathogenetic significance in AML has to be clarified yet.

It is expected that the extension of the present analyses to currently less well-defined AML will identify additional subgroups of AML with clinical relevance based on their gene expression profiles. The feasibility of such an approach has been demonstrated for the first time for diffuse large B-cell lymphoma(45). Alizadeh et al. have subdivided an entity previously considered homogeneous by various pathological methods into two not only new but also prognostically highly relevant subgroups. In two recent studies, gene expression profiling also in breast cancer revealed subgroups significantly differing in their prognosis(46, 47). With regard to AML, this approach may be most promising in AML with normal karyotype. This subgroup cannot be further defined on the cytogenetic level and is characterized by an intermediate prognosis possibly masking poor and favorable subgroups.

In addition, the current data may have major implications with regard to delineating aberrant gene expression pathways underlying the pathogenesis of AML. As has been shown in mantle cell lymphoma and medulloblastoma(48, 49) the extension of our analyses to all subgroups of AML should enable us to define the deregulated genes important for the initiation and the progression of AML. Finally, these analyses will promote the identification of new targets for specific treatment approaches.

EXAMPLE 4 - REFERENCES

1. Bennett, J. M., Catovsky, D., Daniel, M. T., Flandrin, G., Galton, D. A., Gralnick, H. R. & Sultan, C. (1976) *Br. J. Haematol.* **33**, 451-458.

2. Harris, N. L., Jaffe, E. S., Diebold, J., Flandrin, G., Muller Hermelink, H. K., Vardiman, J., Lister, T. A. & Bloomfield, C. D. (1999) *J. Clin. Oncol.* **17**, 3835-3849.

3. Jaffe, E. S., Harris, N. L., Stein, H. & Vardiman, J. W. World Health Organization Classification of Tumours. Pathology and Genetics of Tumours of Haematopoietic and Lymphoid Tissues. 2001. Lyon, IARC Press.

4. Grimwade, D., Walker, H., Oliver, F., Wheatley, K., Harrison, C., Harrison, G., Rees, J., Hann, I., Stevens, R., Burnett, A. *et al.* (1998) *Blood* **92**, 2322-2333.

5. Bloomfield, C. D., Shuma, C., Regal, L., Philip, P. P., Hossfeld, D. K., Hagemeijer, A. M., Garson, O. M., Peterson, B. A., Sakurai, M., Alimena, G. *et al.* (1997) *Cancer* **80**, 2191-2198.

6. Buchner, T., Hiddemann, W., Wormann, B., Loffler, H., Gassmann, W., Haferlach, T., Fonatsch, C., Haase, D., Schoch, C., Hossfeld, D. *et al.* (1999) *Blood* **93**, 4116-4124.

7. Schoch, C., Haferlach, T., Haase, D., Fonatsch, C., Loffler, H., Schlegelberger, B., Staib, P., Sauerland, M. C., Heinecke, A., Buchner, T. *et al.* (2001) *Br. J. Haematol.* **112**, 118-126.

8. Schoch, C., Kern, W., Krawitz, P., Dugas, M., Schnittger, S., Haferlach, T. & Hiddemann, W. (2001) *Blood* **98**, 3500.

9. Pabst, T., Mueller, B. U., Harakawa, N., Schoch, C., Haferlach, T., Behre, G., Hiddemann, W., Zhang, D. E. & Tenen, D. G. (2001) *Nat. Med.* **7**, 444-451.

10. Yuan, Y., Zhou, L., Miyamoto, T., Iwasaki, H., Harakawa, N., Hetherington, C. J., Burel, S. A., Lagasse, E., Weissman, I. L., Akashi, K. *et al.* (2001) *Proc. Natl. Acad. Sci. U. S. A* **98**, 10398-10403.

11. Castilla, L. H., Wijmenga, C., Wang, Q., Stacy, T., Speck, N. A., Eckhaus, M., Marin Padilla, M., Collins, F. S., Wynshaw Boris, A. & Liu, P. P. (1996) *Cell* **87**, 687-696.

12. Brown, D., Kogan, S., Lagasse, E., Weissman, I., Alcalay, M., Pelicci, P. G., Atwater, S. & Bishop, J. M. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94**, 2551-2556.

13. Golub, T. R., Slonim, D. K., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J. P., Coller, H., Loh, M. L., Downing, J. R., Caligiuri, M. A. *et al.* (1999) *Science* **286**, 531-537.

14. Virtaneva, K., Wright, F. A., Tanner, S. M., Yuan, B., Lemon, W. J., Caligiuri, M. A., Bloomfield, C. D., de la Chapelle, A. & Krahe, R. (2001) *Proc. Natl. Acad. Sci. U. S. A* **98**, 1124-1129.

15. Armstrong, S. A., Staunton, J. E., Silverman, L. B., Pieters, R., den Boer, M. L., Minden, M. D., Sallan, S. E., Lander, E. S., Golub, T. R. & Korsmeyer, S. J. (2002) *Nat. Genet.* **30**, 41-47.

16. Haferlach, T., Bennett, J. M., Loffler, H., Gassmann, W., Andersen, J. W., Tuzuner, N., Cassileth, P. A., Fonatsch, C., Schoch, C., Schlegelberger, B. *et al.* (1996) *Leuk. Lymphoma.* **23**, 227-234.

17. Haferlach, T., Winkemann, M., Loffler, H., Schoch, R., Gassmann, W., Fonatsch, C., Schoch, C., Poetsch, M., Weber Matthiesen, K. & Schlegelberger, B. (1996) *Blood* **87**, 2459-2463.

18. Bennett, J. M., Catovsky, D., Daniel, M. T., Flandrin, G., Galton, D. A., Gralnick, H. R. & Sultan, C. (1985) *Ann. Intern. Med.* **103**, 620-625.

19. Dugas, M., Schoch, C., Schnittger, S., Haferlach, T., Danhauser-Riedl, S., Hiddemann, W., Messerer, D. & Uberla, K. (2001) *Leukemia* **15**, 1805-1810.

20. Loffler, H. & Rastetter, J. Atlas of clinical hematology. 1999. Berlin, Springer. Ref Type: Serial (Book,Monograph)

21. Stollmann, B., Fonatsch, C. & Havers, W. (1985) *Br. J. Haematol.* **60**, 183-196.

22. Fonatsch, C., Schaadt, M., Kirchner, H. & Diehl, V. (1980) *Int. J. Cancer* **26**, 749-756.

23. ISCN 1995, Guidelines for Cancer Cytogenetics, Supplement to: An International System for Human Cytogenetic Nomenclature. ed. Mitelman, F. 1995. S. Karger.

24. Evans, P., Jack, A., Short, M., Haynes, A., Shiach, C., Owen, R., Johnson, R. & Morgan, G. J. (1995) *Leukemia* **9**, 1285-1286.

25. Tusher, V. G., Tibshirani, R. & Chu, G. (2001) *Proc. Natl. Acad. Sci. U. S. A* **98**, 5116-5121.

26. Dysvik, B. & Jonassen, I. (2001) *Bioinformatics.* **17**, 369-370.

27. Breiman, L., Friedman, J. H., Olshen, R. A. & Stone, C. J. Classification and regression trees. 1984. Monterey, Wadsworth & Brooks.

28. Quinlan, J. R. Programs for machine learning. 1993. San Mateo, Morgan Kaufmann.

29. Berrar, D., Granzow, M., Dubitzky, W., Lichter, P. & Eils, R. New insights in clinical impact of molecular genetic data by knowledge-driven data mining. 275-281. 2001. Wisconsin, Omnipress. Proceedings of the Second International Conference on Systems Biology.

30. Efron, B. & Tibshirani, R. J. An introduction to the bootstrap. 237-257. 1981. London/New York, Chapman & Hall.

31. Eisen, M. B., Spellman, P. T., Brown, P. O. & Botstein, D. (1998) *Proc. Natl. Acad. Sci. U. S. A* **95**, 14863-14868.

32. Schoch, C., Haase, D., Fonatsch, C., Haferlach, T., Loffler, H., Schlegelberger, B., Hossfeld, D. K., Becher, R., Sauerland, M. C., Heinecke, A. *et al.* (1997) *Br. J. Haematol.* **99**, 605-611.

33. Schnittger, S., Kinkelin, U., Schoch, C., Heinecke, A., Haase, D., Haferlach, T., Buchner, T., Wormann, B., Hiddemann, W. & Griesinger, F. (2000) *Leukemia* **14**, 796-804.

34. Friedman, A. D. (1999) *Leukemia* **13**, 1932-1942.

35. Schnittger, S., Schoch, C., Dugas, M., Kern, W., Staib, P., Wuchter, C., Loffler, H., Sauerland, M. C., Serve, H., Buchner, T. *et al.* (2002) *Blood*, in press.

36. Miyazato, A., Ueno, S., Ohmine, K., Ueda, M., Yoshida, K., Yamashita, Y., Kaneko, T., Mori, M., Kirito, K., Toshima, M. *et al.* (2001) *Blood* **98**, 422-427.

37. Khan, J., Wei, J. S., Ringner, M., Saal, L. H., Ladanyi, M., Westermann, F., Berthold, F., Schwab, M., Antonescu, C. R., Peterson, C. *et al.* (2001) *Nat. Med.* **7**, 673-679.

38. Bittner, M., Meltzer, P., Chen, Y., Jiang, Y., Seftor, E., Hendrix, M., Radmacher, M., Simon, R., Yakhini, Z., Ben Dor, A. *et al.* (2000) *Nature* **406**, 536-540.

39. Ramaswamy, S., Tamayo, P., Rifkin, R., Mukherjee, S., Yeang, C. H., Angelo, M., Ladd, C., Reich, M., Latulippe, E., Mesirov, J. P. *et al.* (2001) *Proc. Natl. Acad. Sci. U. S. A* **98**, 15149-15154.

40. Hedenfalk, I., Duggan, D., Chen, Y., Radmacher, M., Bittner, M., Simon, R., Meltzer, P., Gusterson, B., Esteller, M., Kallioniemi, O. P. *et al.* (2001) *N. Engl. J. Med.* **344**, 539-548.

41. Perou, C. M., Sorlie, T., Eisen, M. B., van de, R. M., Jeffrey, S. S., Rees, C. A., Pollack, J. R., Ross, D. T., Johnsen, H., Akslen, L. A. *et al.* (2000) *Nature* **406**, 747-752.

42. Kohlmann, A., Dugas, M., Schoch, C., Schnittger, S., Mergenthaler, S., Kern, W., Haferlach, T. & Hiddemann, W. (2001) *Blood* **98**, 91a.

43. Miller, J. D., Stacy, T., Liu, P. P. & Speck, N. A. (2001) *Blood* **97**, 2248-2256.

44. Gelmetti, V., Zhang, J., Fanelli, M., Minucci, S., Pelicci, P. G. & Lazar, M. A. (1998) *Mol. Cell Biol.* **18**, 7185-7191.

45. Alizadeh, A. A., Eisen, M. B., Davis, R. E., Ma, C., Lossos, I. S., Rosenwald, A., Boldrick, J. C., Sabet, H., Tran, T., Yu, X. *et al.* (2000) *Nature* **403**, 503-511.

46. Sorlie, T., Perou, C. M., Tibshirani, R., Aas, T., Geisler, S., Johnsen, H., Hastie, T., Eisen, M. B., van de, R. M., Jeffrey, S. S. *et al.* (2001) *Proc. Natl. Acad. Sci. U. S. A* **98**, 10869-10874.

47. West, M., Blanchette, C., Dressman, H., Huang, E., Ishida, S., Spang, R., Zuzan, H., Olson, J. A., Jr., Marks, J. R. & Nevins, J. R. (2001) *Proc. Natl. Acad. Sci. U. S. A* **98**, 11462-11467.

48. Hofmann, W. K., de Vos, S., Tsukasaki, K., Wachsman, W., Pinkus, G. S., Said, J. W. & Koeffler, H. P. (2001) *Blood* **98**, 787-794.

49. MacDonald, T. J., Brown, K. M., LaFleur, B., Peterson, K., Lawlor, C., Chen, Y., Packer, R. J., Cogen, P. & Stephan, D. A. (2001) *Nat. Genet.* **29**, 143-152.

Example 6: Correlation of Protein Expression and Gene Expression in Acute Myeloid Leukemia

INTRODUCTION

The determination of the surface and cytoplasmic expression of characteristic proteins by flow cytometry (FC) is a common method applied to the diagosis and the subclassification of acute myeloid leukemias (AML)[1]. The oligonucleotide microarray analysis (MA) represents a novel technology for the simultaneous detection of the mRNA abundance of large numbers of genes[2,3]. Based on specific gene-expression patterns distinct disease entities have been identified[4-6]. Therefore MA may become of major importance as a diagnostic tool for AML in the near future[7,8]. However, up to now data on the correlation between protein expression levels and mRNA abundance are limited[9-12]. To analyze the relation of protein expression and mRNA abundance in AML we performed 450 individual comparisons of 29 genes in 25 patients with AML at diagnosis analyzed by FC and MA in parallel[13].

METHODS

Samples

Bone marrow samples from highly characterized patients with newly diagnosed and untreated AML were used. Samples had been analyzed by cytomorphology, cytochemistry, cytogenetics and molecular genetics in all cases and were characterized by either of the balanced chromosomal aberrations t(8;21), t(15;17), or inv(16) and the respective molecular and morphologic features[7]. The studies abide by the rules of the local Internal Review Board and the tenets of the revised Helsinki protocol.

Flow cytometry

The studies were performed on cells isolated from bone marrow by Ficoll-Hypaque density gradient centrifugation as described previously[14]. Applying triple-stainings and isotype controls monoclonal antibodies against 29 antigens were used in the following combinations as designed for diagnostic purposes (conjugated with the fluorochromes FITC, PE, and PC-5, respectively): CD34/CD2/CD33, CD7/CD33/CD34, CD34/CD56/CD33, CD11b/CD33/CD34, CD64*/CD4/CD45, CD15*/CD13/CD33, HLA-DR/CD33/CD34, CD34/CD135/CD33,

CD34/CD116/CD33, CD34/NG2/CD33, CD38/CD133**/CD34, CD61/CD14/CD45, CD36/CD235a/CD45, CD34/CD10/CD19, MPO***/LF***/cyCD15, TdT/cyCD22/cyCD3, TdT/cyCD79a/cyCD3. All antibodies were purchased from Immunotech (Marseilles, France), except for: * = Medarex (Annandale, NJ); ** = Milteny Biotech (Bergisch Gladbach, Germany); *** = Caltag (Burlingame, CA). The respective combinations of antibodies were added to $1x10^6$ cells (volume, 100 μl) and incubated for ten minutes at room temperature. The samples were then washed twice in phosphate-buffered saline (PBS) and resuspended in 0.5 ml PBS. FC analysis was performed using a FACSCalibur flow cytometer (Becton Dickinson, San Jose, CA). Analysis of list-mode files was performed by means of the CellQuest Pro Software (Becton Dickinson). Antigen expression was rated positive at a cut-off level of 20% of the cells within the mononuclear gate for membrane proteins and at a cut-off level of 10% for cytoplasmic antigens. Mean fluorescence intensity values were calculated for all events with fluorescence values higher than isotype controls.

Microarray experiments

For microarray analysis the GeneChip® System (Affymetrix, Santa Clara, California) was used. The targets for GeneChip® analysis were prepared according to the current Expression Analysis Technical Manual. Briefly, lysates of the leukemia samples were homogenized (QIAshredder, Qiagen, Hilden, Germany) and total RNA extracted (RNeasy Mini Kit, Qiagen). Normally, 10 μg total RNA isolated from $1x10^7$ cells were used as starting material in the sûbsequent cDNA-synthesis using oligo[$(dT)_{24}$T7promotor]$_{65}$ primer (cDNA Synthesis System, Roche Diagnostics, Mannheim, Germany). The cDNA was purified by phenol:chlorophorm:isoamylalcohol extraction (Ambion, Austin, Texas) and acetate/ethanol precipitated overnight. For detection of the hybridized target nucleic acid biotin-labeled ribonucleotides were incorporated during the *in vitro* transcription (Enzo® BioArray™ HighYield™ RNA Transcript Labeling Kit, ENZO, Farmingdale, USA). After quantification of the purified cRNA (RNeasy Mini Kit, Qiagen), 15 μg were fragmented by alkaline treatment (200 mM Tris-acetate, pH 8.2, 500 mM potassium acetate, 150 mM magnesium acetate) and added to the hybridization cocktail sufficient for 5 hybridizations on standard GeneChip®

microarrays. Before hybridization onto U95Av2, Test3 microarrays (Affymetrix) were chosen for monitoring of labelling efficiency and the integrity of the cRNA. Washing and staining of the probe arrays was performed according to the current protocols (Micro_1v1, EukGE-WS2v4). The Affymetrix software (Microarray Suite, Version 4.0.1) extracted fluorescence intensities from each element on the microarrays as detected by confocal laser scanning according to the manufacturers recommendations. In order to be able to compare different experiments the global
microarray intensities were scaled to a common target intensity. Furthermore, the mRNA abundance of the genes was qualitatively rated as a) present, b) marginal, and c) absent calls, respectively.

Statistics

A total of 29 genes were analyzed in 25 patients with AML. The congruence of positivity and negativity of the expression of the respective genes as determined by FC and MA was analyzed for each gene in each individual patient. Comparisons of microarray intensities were performed by Mann-Whitney *U*-test. Analyses for bivariate correlations of mRNA and protein expression levels were performed by Pearson´s correlation using SPSS, Version 10.0.7.

RESULTS AND DISCUSSION

Twenty-five cases of AML were analyzed in parallel by FC and MA for the expression of 29 genes. Seven had AML M2 with t(8;21), 5 had AML M3 with t(15;17), 7 had AML M3v with t(15;17), and 6 had AML M4Eo with inv(16). A total of 450 comparisons of individual expression data obtained by both methods were performed. Of these, 399 (88.7%) revealed congruent results for protein expression and mRNA abundance (230 cases (51.1%) with positive expression and 169 cases (37.6%) with negative expression, respectively; table 26). In 30 comparisons (6.7%) MA detected positivity for mRNA expression (call: present) while the results of FC indicated negativity. In 21 cases (4.7%) protein expression was demonstrated by FC while no mRNA expression was detected by MA (call: absent).

Focussing on the genes most specific for the diagnosis of AML, i.e. myeloperoxidase, CD13, and CD33, a high correlation between protein expression and mRNA abundance was observed (congruence in 73 of 75 comparisons (97%)). In detail, all cases were rated positive for expression of myeloperoxidase and all but one were positive for both CD13 and CD33, respectively, by both methods. Furthermore, for most other genes essential for the subclassification of AML as well as for the distinction of AML from acute lymphoblastic leukemia and chronic leukemias the results obtained by both methods were always congruent (i.e., for CD10, CD22, CD7, CD133, CD116, CD11b, CD61, CD45, HLA-DR, NG2) or were congruent in the majority (117/140, 84%) of cases (CD79a, CD19, CD2, CD3, CD15, Lactoferrin, CD14, CD235a, CD135, CD34; Table 26). Furthermore, the high correlations between protein expression and mRNA abundance were not limited to congruence in positivity but were significantly correlated also quantitatively. To proof this, the protein expression levels and mRNA abundance were compared by Pearson's correlation in genes expressed in the majority of the analyzed cases. These comparisons revealed significant correlations for the fluorescence intensities as assessed by FC and MA for CD13 ($p=0.001$), CD33 ($p=0.034$), CD34 ($p=0.003$), CD45 ($p=0.015$), CD15 ($p=0.016$), and CD7 ($p=0.033$) and thus further underline the high coherence of expression patterns for both protein and mRNA (figure 19).

Thirty comparisons displayed mRNA expression and no protein expression. Due to the ongoing process of maturation (CD14, CD15) and due to the cross-lineage expression of the genes (CD3, CD19) the levels of mRNA abundance may have been to low to result in detectable protein expression levels using the described cut-off levels of 20% and 10%, respectively. This suggestion is supported by a quantitative analysis of mRNA expression data which shows relatively low albeit positive levels for the respective cases and genes (mean average fluorescence intensity, 46.7 ± 54.5 in cases positive for CD14, CD15, CD3, or CD19 versus 389.4 ± 831.0 in all positive cases, Mann-Whitney *U*-test: $p<0.001$) while at the same time protein expression amounts to a mean of $5\pm4\%$.

Twenty-one comparisons displayed positivity by FC and negativity of MA, which comprize 4.7% of all individual comparisons performed. These discrepancies most probably are due to: a) erythrocytic debris positive for CD36 interfering with the

acquisition of CD36 negative cells during flow cytometric analysis; b) differences between both methods in the selected DNA sequences and antigen epitopes, respectively, detected (i.e. CD38, CD4, CD56); and c) differences in the stability of mRNA and protein of the respective genes.

Overall, these results demonstrate for the first time that there is a significant correlation between protein expression and gene expression in AML and that the antigens so far identified essential for the diagnosis and subclassification of AML by flow cytometry may represent additional candidate genes when using MA as a diagnostic tool for molecular cancer class prediction[15,16]. Furthermore, it is anticipated that the present analyses represent a prime example and will be reproduced for a variety of other entities like lymphoid malignancies. Due to their high potential to assess the expression patterns of high numbers of genes and due to their excellent reproducibility features microarrays are a promising future diagnostic tool. As a consequence, they may replace the more time and resource consuming diagnostic methods currently used for diagnosing leukemias like cytomorphology, cytogenetics, and FC.

REFERENCES for Example 5

1. Weir, E. G. and Borowitz, M. J. Flow cytometry in the diagnosis of acute leukemia. Semin.Hematol., *38*: 124-138, 2001.
2. Liotta, L. and Petricoin, E. Molecular profiling of human cancer. Nat.Rev.Genet., *1*: 48-56, 2000.
3. Kohlmann, A., Dugas, M., Schoch, C., Schnittger, S., Mergenthaler, S., Kern, W., Haferlach, T., and Hiddemann, W. Gene expression profiles of distinct AML subtypes in comparison to normal bone marrow. Blood, *98*: 91a, 2001.
4. Alizadeh, A. A., Eisen, M. B., Davis, R. E., Ma, C., Lossos, I. S., Rosenwald, A., Boldrick, J. C., Sabet, H., Tran, T., Yu, X., Powell, J. I., Yang, L., Marti, G. E., Moore, T., Hudson, J., Jr., Lu, L., Lewis, D. B., Tibshirani, R., Sherlock, G., Chan, W. C., Greiner, T. C., Weisenburger, D. D., Armitage, J. O., Warnke, R., Staudt, L. M., and . Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature, *403*: 503-511, 2000.

5. Hedenfalk, I., Duggan, D., Chen, Y., Radmacher, M., Bittner, M., Simon, R., Meltzer, P., Gusterson, B., Esteller, M., Kallioniemi, O. P., Wilfond, B., Borg, A., and Trent, J. Gene-expression profiles in hereditary breast cancer. N.Engl.J.Med., *344*: 539-548, 2001.

6. Khan, J., Wei, J. S., Ringner, M., Saal, L. H., Ladanyi, M., Westermann, F., Berthold, F., Schwab, M., Antonescu, C. R., Peterson, C., and Meltzer, P. S. Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks. Nat.Med., *7*: 673-679, 2001.

7. Schoch, C., Kohlmann, A., Schnittger, S., Mergenthaler, S., Dugas, M., Kern, W., Hiddemann, W., and Haferlach, T. Specific abnormalities on the genomic level result in a distinct gene expression pattern detected by oligonucleotide microarrays: an analysis of 25 patients with AML M2/t(8;21), AML M3/M3v/t(15;17) and AML M4eo/inv(16). Blood, *98*: 92a, 2002.

8. Armstrong, S. A., Staunton, J. E., Silverman, L. B., Pieters, R., den Boer, M. L., Minden, M. D., Sallan, S. E., Lander, E. S., Golub, T. R., and Korsmeyer, S. J. MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia. Nat.Genet., *30*: 41-47, 2002.

9. Ahram, M., Best, C. J., Flaig, M. J., Gillespie, J. W., Leiva, I. M., Chuaqui, R. F., Zhou, G., Shu, H., Duray, P. H., Linehan, W. M., Raffeld, M., Ornstein, D. K., Zhao, Y., Petricoin, E. F., III, and Emmert-Buck, M. R. Proteomic analysis of human prostate cancer. Mol.Carcinog., *33*: 9-15, 2002.

10. Serrano, J., Roman, J., Jimenez, A., Castillejo, J. A., Navarro, J. A., Sanchez, J., Garcia-Castellanos, J. M., Martin, C., Maldonado, J., and Torres, A. Genetic, phenotypic and clinical features of acute lymphoblastic leukemias expressing myeloperoxidase mRNA detected by RT-PCR. Leukemia, *13*: 175-180, 1999.

11. Zhan, F., Hardin, J., Kordsmeier, B., Bumm, K., Zheng, M., Tian, E., Sanderson, R., Yang, Y., Wilson, C., Zangari, M., Anaissie, E., Morris, C., Muwalla, F., van Rhee, F., Fassas, A., Crowley, J., Tricot, G., Barlogie, B., and Shaughnessy, J., Jr. Global gene expression profiling of multiple myeloma, monoclonal gammopathy of undetermined significance, and normal bone marrow plasma cells. Blood, *99*: 1745-1757, 2002.

12. Falini, B. and Mason, D. Y. Proteins encoded by genes involved in chromosomal alterations in lymphoma and leukemia: clinical value of their detection by immunocytochemistry. Blood, *99*: 409-426, 2002.

13. Kern, W., Kohlmann, A., Wuchter, C., Schnittger, S., Schoch, C., Mergenthaler, S., Ratei, R., Ludwig, W. D., Hiddemann, W., and Haferlach, T. Correlation of gene expression and protein expression in acute myeloid leukemia as assessed by microarray analysis and flow cytometry. Blood, *98*: 92a, 2001.

14. Wuchter, C., Harbott, J., Schoch, C., Schnittger, S., Borkhardt, A., Karawajew, L., Ratei, R., Ruppert, V., Haferlach, T., Creutzig, U., Dorken, B., and Ludwig, W. D. Detection of acute leukemia cells with mixed lineage leukemia (MLL) gene rearrangements by flow cytometry using monoclonal antibody 7.1. Leukemia, *14*: 1232-1238, 2000.

15. Golub, T. R., Slonim, D. K., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J. P., Coller, H., Loh, M. L., Downing, J. R., Caligiuri, M. A., Bloomfield, C. D., and Lander, E. S. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science, *286*: 531-537, 1999.

16. Ramaswamy, S., Tamayo, P., Rifkin, R., Mukherjee, S., Yeang, C. H., Angelo, M., Ladd, C., Reich, M., Latulippe, E., Mesirov, J. P., Poggio, T., Gerald, W., Loda, M., Lander, E. S., and Golub, T. R. Multiclass cancer diagnosis using tumor gene expression signatures. Proc.Natl.Acad.Sci.U.S.A, *98*: 15149-15154, 2001.

Example 6: Gene Expression Profiles of Distinct Cytogenetic AML Subtypes as Defined by the New WHO Classification: A Study of 45 Patients

Example 6: Introduction

Since their introduction, microarrays have been promising tools for basic research. With regard to leukemia, the pivotal discrimination of unselected acute lymphoblastic (ALL), and acute myeloid leukemia (AML) samples based on their gene expression signatures inspired numerous studies (Golub et al., 1999). We performed gene expression analyses to designate candidate genes for discriminating specific AML samples from normal bone marrow (BM) of healthy volunteers. With regard to the classification of hematological malignancies according to the WHO, distinct AML subtypes have been established based on genetic abnormalities of the leukemic blasts. Here, we demonstrate gene expression analyses of 8 healthy BM donors and 45 leukemia patients representing four cytogenetic subtypes of AML: t(8;21)(q22;q22), inv(16)(p13q22), t(15;17)(q22;q12), and t(11q23)/MLL. Combining different approaches for data analysis a minimal set of genes was identified to designate a reliable class prediction model. Based on the expression pattern of 39 genes, cytogenetically defined AML subtypes could accurately be predicted and separated from healthy BM. Taken together, gene expression signatures of AML cases with recurrent genetic abnormalities demonstrate a very close correlation between genotype and gene expression. Therefore, introducing a set of candidate genes, expression profiling may serve for diagnosis of AML subtypes defined by the new WHO classification.

Example 6 Material and Methods

We analyzed BM aspirates from 8 healthy volunteers and the following 45 untreated AML patients:

- t(8;21)(q22;q22)/AML M2 (n=9),
- t(15;17)(q22;q12)/AML M3/M3v (n=16),

- inv(16)(p13q22)/AML M4eo (n=10), and
- t(11q23)/MLL-aberrations (n=10)

Example 6- Microarray experiments. Gene expression analyses were performed from cells remaining from the diagnostic sample. They had immediately been lysed, frozen and were stored at -80°C from 1 to 34 months until preparation for gene expression profiling. The targets for U95Av2 microarrays were prepared according to current protocols (Affymetrix). Before expression profiling, Test3 Probe Arrays were chosen for monitoring the integrity of the cRNA.

Example 6 - Results I: Characterization of leukemia samples

AML samples were thoroughly characterized by a combination of cytomorphology, cytogenetics, FISH, RT-PCR and quantitative real-time PCR (Fig. 20). All patients showed the above mentioned balanced abnormalities as the sole karyotype change. Using FISH analysis, more than 90% of cells demonstrated the specific signal constellation. The respective fusion transcripts AML1-ETO in t(8;21), CBF□-MYH11 in inv(16), PML-RAR□ in t(15;17) and various MLL-fusion partners in t(11q23) were detected by PCR techniques in all samples. These subtypes are specifically associated with five cytomorphological subtypes according to FAB classification: inv(16)(p13q22)/AML M4eo, t(8;21)(q22;q22)/AML M2, t(15;17)(q22;q12)/AML M3/M3v, and t(11q23)/MLL in FAB M5a/b, respectively. AML subtypes M3 and M3v both carry the same chromosome aberration but differ in morphological and clinical aspects.

Example 6 - Results II: Class separation

For data analysis we combined different approaches. First, a reduced subset of 200 genes obtained by permutation-based neighborhood analysis (SAM, Tusher et al., 2001) was visualized for corresponding clusters using principal component analysis (J-Express, Dysvik et al., 2001)(Fig.21). Samples from healthy donors cluster into a distinct group, likewise all AML samples demonstrate homogenity by forming a second cluster.

Example 6 - Results III: Class prediction

Next, we adapted the signal-to-noise/weighted voting algorithm (Golub et al., 1999) to identify discriminative genes. A minimal set of 39 genes, which provided both optimal classification accuracy and highest prediction strength, was selected to avoid overfitting. The significance of each gene was tested by permutation-based neighborhood analysis. The robustness of the classifier was assessed by leave-one-out crossvalidation. These expression signatures were sufficient to distinguish AML samples with high accuracies from normal bone marrow and to predict the recurrent chromosome aberration, respectively (Table 27, Fig. 22). Table 28a shows for which comparison a gene was important including its statistical significance.

A set of 39 genes is sufficient for class prediction. *Accuracy* denotes the rate of correctly classified test samples. P(g,c) indicates the signal-to-noise ratio of gene x: $S_x = (\mu_1-\mu_2)/(\sigma_1+\sigma_2)$, where μ_k and σ_k denote the mean expression and standard deviation of gene x in group k. As calculated from pairwise comparisons (class A vs. B), positive P(g,c) values indicate a higher gene expression in class A, negative P(g,c) values a higher gene expression in class B, respectively. HGNC symbols are given in column 1.

All leukemia samples could accurately be assigned to their corresponding cytogenetic subtype with 100% accuracies. To illustrate these results, a hierarchical clustering is shown (Fig. 23).

Example 6 – Conclusions

- The expression pattern of 39 genes allowed precise class assignments of four cytogenetically defined AML subtypes according to the WHO classification of hematological malignancies, and normal BM, respectively.
- Thus, we introduce candidate genes suitable for diagnosis of AML subgroups based on gene expression profiling.
- Potentially, gene expression patterns will allow the additional subclassification of AML, especially in subtypes with no specific cytogenetic markers (e.g. normal karyotype).

Example 7: Gene Expression Profiles of Distinct Leukemia Types and Subtypes: A Study of 280 Patients using high-density microarrays

Example 7: Introduction

Here, we demonstrate gene expression analyses of 9 healthy BM donors and 271 leukemia patients representing:

AML: 4 distinct cytogenetic subtypes t(8;21)(q22;q22) (AML t(8;21)), inv(16)(p13q22) (AML inv(16)), t(15;17)(q22;q12) (AML t(15;17)), and t(11q23)/MLL (AML MLL). In addition, we analyzed AML samples characterized by normal karyotypes (AML normal), complex aberrant karyotypes (AML complex), trisomy 8 as sole aberration (AML +8), and other chromosomal changes (AML other).

ALL: 3 distinct genetically defined subtypes: t(4;11)(q21;q23) (ALL t(4;11)), t(8;14)(q24;q32) (ALL t(8;14)), t(9;22)(q34;q11) (ALL Ph) and 2 subtypes defined by their immunophenotype: ALL of the B-lineage not carrying the t(9;22) (ALL B not Ph) and T-ALL (T-ALL)

CLL: 5 genetically defined subtypes: trisomy 12 (tri 12), deletion 11q (11q-), deletion 13q (13q-), deletion 17p (17p-) and none of these aberrations (normal)

CML (CML) without any further subdivison and

Normal bone marrow from healthy volunteers (normal BM).

We used the Affymetrix oligonucleotide microarray technology (GeneChip® Instrument System) to obtain gene expression profiles of each individual clinical sample of interest. The commercially available HG-U133 probe arrays gave information about the relative mRNA abundance of about 33,000 human genes which are represented on these high-density DNA-oligonucleotide microarrays.

Chip Information (as provided by manufacturer):

The GeneChip® Human Genome U133 Set (HG-U133A and HG-U133B) is comprised of two microarrays containing over 1,000,000 unique oligonucleotide features covering more than 39,000 transcript variants, which in turn represent greater than 33,000 of the best characterized human genes. This powerful set allows to reproducibly examine the quantitative and qualitative expression of most genes in the human genome, and was designed using the recently published and publicly available draft of the human genome sequence. Sequences used in the design of the array were selected from GenBank, dbEST, and RefSeq. Sequence clusters were created from Build 133 of UniGene (April 20, 2001) and refined by analysis and comparison with a number of other publicly available databases including the Washington University EST trace repository and the University of California, Santa Cruz golden-path human genome database (April 2001 release). In addition, ESTs were analyzed for untrimmed low-quality sequence information, correct orientation, false priming, false clustering, alternative splicing and alternative polyadenylation.

Combining different approaches for data analysis, a set of genes was identified to designate a reliable class prediction model. Based on the expression pattern of those genes, defined leukemia types and subtypes could accurately be predicted and separated from healthy BM. Taken together, gene expression signatures demonstrate a very close correlation between genotype and gene expression. Therefore, introducing a set of candidate genes, measurements of mRNA abundancies by gene expression profiling serves for diagnosis of leukemia types and subtypes.

Example 7 Material and Methods

We analyzed BM aspirates from 9 healthy volunteers and the following 280 leukemia patients:

Acute myeloid leukemia (AML)

t(8;21)(q22;q22)/AML M2 (n=13),

t(15;17)(q22;q12)/AML M3/M3v (n=20),

inv(16)(p13q22)/AML M4eo (n=12),

t(11q23)/MLL-aberrations (n=15)

trisomy 8 (n=10)

normal karyotype (n=62)

complex aberrant karyotype (n=36)

other aberrations (n=5)

Acute lymphoblastic leukemia (ALL)

t(4;11)(q21;q23) (n=9)

t(8;14)(q24;q32) (n=4)

t(9;22)(q34;q11) (ALL Ph) (n=15)

ALL B lineage without t(9;22) (ALL B not Ph) (n=9)

T-ALL (n=9)

Chronic lymphocytic leukemia (CLL)

trisomy 12 (tri 12) (n=5)

deletion 11q (11q-) (n=4)

deletion 13q (13q-) (n=10)

deletion 17p (17p-) (n=4)

none of these aberrations (normal) (n=9)

Chronic myeloid leukemia (n=14)

Normal bone marrow (normal BM) (n=9)

Example 7 - Results I: Characterization of leukemia samples

We selected bone marrow (BM) samples from 271 leukemia patients at diagnosis representing 18 different disease entities or subentities and from 9 healthy volunteers, respectively. All cases were sent for reference diagnostics to our laboratory, registered in our leukemia database and were treated within prospective randomized multi-center trials. The studies abide by the rules of the local internal review board and the tenets of the revised Helsinki protocol. Samples were received either locally or by overnight mail. Diagnosis was performed by an individual combination of cytomorphology, cytogenetics, FISH, immunophenotyping and molecular genetics. Mononuclear cells were isolated by a Ficoll gradient, lysed, frozen and were stored at -80°C from one to 34 months until sample preparation for gene expression analysis. All leukemia samples were thoroughly characterized by a individual combination of cytomorphology, cytogenetics, immunophenotyping, fluorescence in situ hybridisation (FISH), polymerase chain reaction based methods both qualitative RT-PCR and quantitative real-time PCR. Using FISH analysis, more than 90% of cells demonstrated the specific signal constellation. The respective fusion transcripts BCR-ABL in t(9;22) positive CML (Schoch et al. 2002a) and in t(9;22) positive ALL, AML1-ETO in AML with t(8;21), CBFbeta-MYH11 in AML with inv(16), PML-RARalpha in AML with t(15;17) (Schoch et al. 2002b) and various MLL-fusion partners in both AML and ALL with t(11q23) were detected by FISH and PCR techniques in all samples.

In t(8;14) positive ALL the IGH-C-MYC rearrangement was confirmed by FISH. In all cases with AML and complex aberrant karyotype 24 color FISH was performed in addition to chromosome banding analysis (Schoch et al. 2002c).

Genetic subtyping of CLL was carried out using interphase FISH with the following probes (Buhmann et al. 2002):

- for the detection of trisomy 12 a centromere specific probe for chromosome 12
- for the detection of 11q deletions probes for the ATM as well as for the RDX gene
- for the detection of 13q deletions probes for the retinoblastoma gene (Rb), and the anonymous loci D13S25 and D13S319
- for the detection of 17p deletion a probe for the p53 gene
- cases with none of the above mentioned aberrations were assigned to the group normal

References:.

Buhmann R, Kurzeder C, Rehklau J, Westhaus D, Bursch S, Hiddemann W, Haferlach T, Hallek M, Schoch C.

CD40L stimulation enhances the ability of conventional metaphase cytogenetics to detect chromosome aberrations in B-cell chronic lymphocytic leukaemia cells.

Br J Haematol 2002 Sep;118(4):968-75

Schoch C, Schnittger S, Kern W, Lengfelder E, Löffler H, Hiddemann W, Haferlach T.

Rapid diagnostic approach to PML-RARalpha-positive acute promyelocytic leukemia.
Hematol J 2002a;3(5):259-63

Schoch C, Schnittger S, Bursch S, Gerstner D, Hochhaus A, Berger U, Hehlmann R, Hiddemann W, Haferlach T.

Comparison of chromosome banding analysis, interphase- and hypermetaphase-FISH, qualitative and quantitative PCR for diagnosis and for follow-up in chronic myeloid leukemia: a study on 350 cases, Leukemia 2002b Jan;16(1):53-9

Schoch C, Haferlach T, Bursch S, Gerstner D, Schnittger S, Dugas M, Kern W, Löffler H, Hiddemann W.

Loss of genetic material is more common than gain in acute myeloid leukemia with complex aberrant karyotype: A detailed analysis of 125 cases using conventional chromosome analysis and fluorescence in situ hybridization including 24-color FISH.
Genes Chromosomes Cancer 2002 Sep;35(1):20-9

Example 7 - Results II: Sample preparation and microarray hybridisation

Microarray analyses were performed utilising the GeneChip® System (Affymetrix, Santa Clara, USA). The targets for GeneChip® analyses were prepared according to the current Expression Analysis Technical Manual. Briefly, lysates of the leukemia samples were homogenised (QIAshredder, Qiagen, Hilden, Germany) and total RNA extracted (RNeasy Mini Kit, Qiagen). Normally, 5 μg total RNA isolated from 1×10^7 cells were used as starting material in the subsequent cDNA-synthesis using oligo[$(dT)_{24}$T7promotor]$_{65}$ primer (cDNA Synthesis System, Roche Applied Science, Mannheim, Germany). The cDNA was purified by phenol:chloroform:isoamyl alcohol (25:24:1) extraction (Ambion, Austin, USA) and acetate/ethanol precipitated over night. For detection of the hybridised target nucleic acid biotin-labeled ribonucleotides were incorporated during the in vitro transcription (Enzo® BioArray™ HighYield™ RNA Transcript Labeling Kit, ENZO, Farmingdale, USA). After quantification of the purified cRNA (RNeasy Mini Kit, Qiagen), 15 μg labeled cRNA were fragmented by alkaline treatment (200 mM Tris-acetate, pH 8.2, 500 mM potassium acetate, 150 mM magnesium acetate) and added to the hybridisation cocktail sufficient for 5 hybridisations on standard

format GeneChip® microarrays. Before hybridisation to HG-U133 microarrays, Test3 microarrays (Affymetrix) were chosen in some cases for monitoring the integrity of the cRNA. Washing and staining of the probe arrays was performed according to the current protocols of the manufacturer (Fluidics Station, Micro_1v1, EukGE-WS2v4). The Affymetrix software (Microarray Suite, Version 5.0) extracted fluorescence intensities from each feature on the microarrays as detected by confocal laser scanning according to the manufacturers recommendations. Some of the hybridization cocktails had previously been hybridized to U95Av2 arrays. Hybridization cocktails can be used for up to 5 distinct array analyses.

All hybridisation cocktails demonstrated high quality cRNA characteristics. We considered both low 3'/5' ratio (e.g., lower than about 3) of housekeeping controls and the total number of present called genes (> about 30% on U133A), along with the average signal intensity of a present called gene. Expression profiles which fulfilled all quality control criteria were selected for subsequent supervised selection of informative genes.

Example 7 - Results III: Statistical Analyses

For data analysis we combined different approaches. First, the expression data was preprocessed. Raw expression intensities were scaled using the Affymetrix Microarray Suite software scaling parameter (target intensity: 5000). This preprocessing is based on a mask file which compares expression intensities of a set of 100 genes which code for ubiquitous housekeeping cellular proteins. This set of genes for normalisation of expression intensities is represented on both U133A and U133B arrays. The step of data preprocessing assures that array experiments can be compared properly using further statistical algorithms and methods. Subsequently, the data was analyzed according to two different established methods from as described below. The results from the two analyses were systematically compared to validate the list of differentially expressed genes.

1. Selection of differentially expressed genes

a) Analysis according to example 3.

The top 20 differentially expressed genes were calculated for all disease entities and normal bone marrow, respectively, as described in example 3. Expression data were analyzed in order to select a minimal set of discriminative genes, which provides, as described hereinabove (Example 3), maximum classification accuracy in leave-one-out-crossvalidation.

One-versus-all (OVA) and all-pairs comparisons (AP) were systematically applied. Genes were ranked according to signal-to-noise ratio (STN). For each OVA and AP comparison a set of discriminative genes is disclosed in tables 29, 32, 35, 38 and 41 whereby the gene names can be found in tables 43a,b. The most discriminative and informative genes are marked by asterisks in tables 29, 32, 35, 38 and 41. Classification accuracy was estimated by means of leave-one-out-crossvalidation and weighted voting.

References:

Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, Coller H, Loh ML, Downing JR, Caligiuri MA, Bloomfield CD, Lander ES. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science 1999; 286(5439):531-7

Pomeroy SL, Tamayo P, Gaasenbeek M, Sturla LM, Angelo M, McLaughlin ME, KimJY, Goumnerova LC, Black PM, Lau C, Allen JC, Zagzag D, Olson JM, Curran T,Wetmore C, Biegel JA, Poggio T, Mukherjee S, Rifkin R, Califano A, StolovitzkyG, Louis DN, Mesirov JP, Lander ES, Golub TR. Prediction of central nervous system embryonal tumour outcome based on geneexpression. Nature 2002; 415(6870):436-42.

2. Estimation of classification accuracy

A set of 20 top-ranked genes, which provided both optimal classification accuracy and highest prediction strength, was selected to avoid overfitting. The significance

of each gene was tested by permutation-based neighborhood analysis. The robustness of the classifier was assessed by leave-one-out crossvalidation. These expression signatures were sufficient to distinguish leukemia samples with high accuracies from normal bone marrow and also to predict the recurrent chromosome aberration, respectively (Tables 29, 32, 35, 38, 41). *Accuracy* denotes the rate of correctly classified test samples. P(g,c) indicates the signal-to-noise ratio of gene x: $S_x = (\mu_1-\mu_2)/(\sigma_1+\sigma_2)$, where μ_k and σ_k denote the mean expression and standard deviation of gene x in group k. As calculated from pairwise comparisons (class A vs. B), positive P(g,c) values indicate a higher gene expression in class A, negative P(g,c) values a higher gene expression in class B, respectively.

b) Analysis according to Westfall & Young the same data set was analysed according to Westfall & Young to identify significantly differentially expressed genes with adjustment of the p-values for multiple testing.

Step-down maxT and minP multiple testing procedures were applied, which compute permutation adjusted p-values for the step-down maxT and minP multiple testing procedures, which provide strong control of the family-wise Type I error rate (FWER). The multitest package (version 1.0) from Bioconductor was applied, which is based on the R statistical language. These methods outperform other methods (see Dudoit, JASA 2002).

References:

Westfall PH, Young SS (1993) Resampling-based multiple testing: Examples and methods for p-value adjustment. John Wiley & Sons. ISBN 0-471-55761-7

Dudoit S, Fridlyand J, Speed TP.

Comparison of Discrimination Methods for the Classification of Tumors Using Gene Expression Data. JASA 2002; 97:77-87

Package multtest (version 1.0)

from Bioconductor http://www.bioconductor.org

R statistical language: http://www.r-project.org/

c) Comparison of gene lists

The list of differentially expressed genes obtained from 1a) and 1b) were systematically compared using PERL scripts in order to identify genes that occurred in both list, versus genes occurring in one list only.

Expression intensities (expression levels) derived from the above-mentioned MicroArray Suite program were plotted as bar graphs showing gene expression profiles using a Perl script (Figures 24 to 464).

References:

PERL: http://www.perl.com

Sensitivities for the detection of leukemia types and subtypes were calculated as the number of positive samples predicted divided by the number of true positives.

Specificities for the detection of leukemia types and subtypes were calculated as the number of negative samples predicted divided by the number of true negatives.

Example 7 - Results IV: Analysis of 14 leukemia subtypes and normal bone marrow

Here we analyzed in total 14 distinct leukemia types and subtypes as well a cohort of healthy volunteers for normal bone marrow characteristics. We applied the described two different statistical methods for identification of genes which allow accurate class assignments to the respective groups.

ALL t(4;11) (n=9)

ALL t(8;14) (n=4)

ALL B not Ph (n=9)

ALL Ph (n=15)

T-ALL (n=9)

AML +8 (n=10)

AML complex (n=36)

AML normal (n=62)

AML t(8;21) (n=13)

AML t(15;17) (n=20)

AML inv(16) (n=12)

AML MLL (n=15)

CLL (n=32)

CML (n=14)

normal BM (n=9)

total: 269 samples

First, expression data were analyzed according to example 3, as described hereinabove.

A set of 20 top-ranked genes, which provided both optimal classification accuracy and highest prediction strength for all pairwise (all pairs) and one-versus-all comparisons is given as table 29. Within this set of genes, optimal classification accuracy can be obtained with genes marked by asterisks. Gene expression intensities, plotted as bar graphs are given in Figures 24 to 188. Genes are depicted as unique Affymetrix identifier (for example 201497_x_at) and, where available, approved HGNC symbols (HUGO Gene Nomenclature Committee). More detailed, the complete annotation and sequence information about this set of genes is listed in tables 43a,b.

In total 269 cases with leukemia or normal bone marrow (BM) were analyzed. 248 of 269 (92.2%) cases were assigned to the correct leukemia type in all pairwise comparisons (table 28 b). The sensitivity indicated for each subgroup indicates the percentage of cases of the specific subgroup identified correctly in all pairwise comparisons (range 60% to 100%). The specificity indicates for each subgroup the percentage of correct assignments to this subgroup (range 85.3% to 100%).

In total 3766 individual assignments of leukemia and normal bone marrow were analyzed. 3745 of 3766 assignments (99.4%) were correct (table 28c). The sensitivity indicated for each subgroup indicates the percentage of correct assignments for cases of the specific subgroup in pairwise comparisons. (range 97.1% to 100%). The specificity indicates for each subgroup the percentage of correct assignments to this subgroup (range 98.4% to 100%).

In a second approach significant genes were identified according to Westfall & Young. Table 30 represents all genes found to be significant after p-value adjustment. Genes are depicted as unique Affymetrix identifier (for example 201497_x_at) and, where available, approved HGNC symbols (HUGO Gene Nomenclature Committee). More detailed, the complete annotation and sequence information about this set of genes is listed in table 43a,b.

Furthermore, we provide information about genes which were found to be rated significant independently by both methodologies (Table 30). Top-significant genes

according to the method of example 3 are marked by asterisks. Genes which were included in any of the top-20 lists are marked by positive signs.

In addition, selected gene profiles were chosen to demonstrate their capability of discriminating different leukemia types, subtypes and normal bone marrow, respectively. Gene expression profiles were generated by means of PERL-programs, evaluated and plotted as bar graphs. Each of the analyzed groups are accordingly outlined. The following genes were selected and are given as Figures 189 to 233:

GeneID	gene symbol	feature
201162_at	IGFBP7	CLL low
201163_s_at	IGFBP7	CLL low
201362_at	NS1-BP	CML high
201496_x_at	MYH11	AML inv(16) high
201497_x_at	MYH11	AML inv(16) high
201998_at	SIAT1	CLL high
202095_s_at	BIRC5	CLL low
203074_at	ANXA8	AML t(15;17) high
204150_at	STAB1	AML t(15;17) high
204511_at	KIAA0793	CLL high
205528_s_at	CBFA2T1	AML t(8;21) high
205529_s_at	CBFA2T1	AML t(8;21) high

205805_s_at	ROR1	CLL high
206940_s_at	POU4F1	AML t(8;21) high
207819_s_at	ABCB4	CLL high
208091_s_at	DKFZP564K0822	CLL high
208456_s_at	RRAS2	CLL high
209061_at	NCOA3	CLL high
209101_at	CTGF	ALL t(4;11) high, ALL Ph high, T-ALL high
209374_s_at	IGHM	CLL high
209616_s_at	CES1	AML MLL high
210997_at	HGF	AML t(15;17) high
212285_s_at	AGRN	AML t(15;17) high
213539_at	CD3D	T-ALL high
214450_at	CTSW	AML t(15;17) high
215925_s_at		ALL t(4;11) high
218223_s_at	LOC51177	CML low
222166_at		AML +8 high
224520_s_at	MGC13168	ALL t(8;14) high
224794_s_at	LOC51148	AML t(15;17) high
225660_at	SEMA6A	ALL B not Ph high, ALL Ph high

226496_at	Homo sapiens, Similar to hypothetical protein FLJ22611, clone MGC:24716 IMAGE:4277726, mRNA, complete cds	ALL high, CLL high
228827_at	Homo sapiens clone 25023 mRNA sequence	AML t(8;21) high
228904_at	ESTs	AML normal high, AML +8 high, AML complex high
236301_at	Homo sapiens, clone IMAGE:3866403, mRNA	CLL high
236892_s_at	HOXB6	AML normal high, AML +8 high, AML complex high
239214_at	ESTs	ALL t(4;11) high
239393_at	ESTs	ALL t(4;11) high
239791_at	HOXB6	AML normal high, AML +8 high
240581_at	ESTs	ALL t(4;11) high
241464_s_at	ESTs	AML MLL high, AML normal high, AML +8 high, AML complex high
241525_at	ESTs	AML inv(16) high
243362_s_at	LEF1	ALL high, CLL high
36566_at	CTNS	T-ALL low
38487_at	FLJ12442	AML t(15;17) high

Generally, chromosomal aberrations are strongly associated with morphological characteristics. However, there are two chromosomal aberrations which are observed in both myeloid and lymphatic neoplasms, i.e. t(11q23)/MLL and the t(9;22). The t(9;22) occurs in ALL (ALL Ph) and CML, and t(11q23)/MLL is observed in ALL (ALL t(4;11)) and AML (AML MLL), respectively. Analysing gene expression signatures of both t(9;22) positive ALL and CML we identified genes, which allowed correct lineage assignments (table 29). In addition, our results indicate that the distinct expression signatures are also sufficient for correct assignments of the t(11q23)/MLL positive leukemias either to ALL or to AML (table 29). Thus, in both scenarios lineage assignment (lymphoid or myeloid), and even subtype classification can be accomplished based on the methods and markers described herein, despite of the fact that e.g., in the above-noted t(11q23) and t(9;22) chromosomal aberrations, the same chromosomal aberration is associated with different kinds of leukemia.

Example 7 - Results V: Analysis of 5 ALL subtypes defined by genetics and Immunophenotype

Here we analyzed in 5 distinct ALL subtypes. We applied the described two different statistical methods for identification of genes which allow accurate class assignments to the respective groups.

ALL t(4;11)	(n=9)
ALL t(8;14)	(n=4)
ALL B not Ph	(n=9)
ALL Ph	(n=15)
T-ALL	(n=9)

First, expression data were analyzed according to example 3, as described hereinabove.

A set of 20 top-ranked genes, which provided both optimal classification accuracy and highest prediction strength for all pairwise (all pairs) and one-versus-all comparisons is given in table 32. Within this set of genes, optimal classification accuracy can be obtained with genes marked by asterisks. Gene expression intensities, plotted as bar graphs are given in Figures 234 to 252. Genes are depicted as unique Affymetrix identifier (for example 201497_x_at) and, where available, approved HGNC symbols (HUGO Gene Nomenclature Committee). More detailed, the complete annotation and sequence information about this set of genes is listed in table 43a,b.

In total 46 cases of ALL were analyzed. 44 of 46 cases (95.7%) were assigned to the correct ALL subtype in all pairwise comparisons (table 31a). The sensitivity indicated for each subgroup indicates the percentage of cases of the specific subgroup identified correctly in all pairwise comparisons (range 88.9% to 100%). The specificity indicates for each subgroup the percentage of correct assignments to this subgroup (range 88.9% to 100%).

In total 184 individual assignments of ALL were analyzed. 182 of 184 assignments (98.9%) were correct (table 31b). The sensitivity indicated for each subgroup indicates the percentage of correct assignments for cases of the specific subgroup in pairwise comparisons. (range 97.2% to 100%). The specificity indicates for each subgroup the percentage of correct assignments to this subgroup (range 97.2% to 100%).

In a second approach significant genes were identified according to Westfall & Young. Table 33 represents all genes found to be significant after p-value adjustment. Genes are depicted as unique Affymetrix identifier (for example 201497_x_at) and, where available, approved HGNC symbols (HUGO Gene Nomenclature Committee). More detailed, the complete annotation and sequence information about this set of genes is listed in table 43a,b.

Furthermore, we provide information about genes which were found to be rated significant independently by both methodologies (Table 33). Top-significant genes according to the method of example 3 hereinabove are marked by asterisks. Genes which were included in any of the top-20 lists are marked by positive signs.

In addition, selected gene profiles were chosen to demonstrate their capability of discriminating different leukemia types, subtypes and normal bone marrow, respectively. Gene expression profiles were generated by means of PERL-programs, evaluated and plotted as bar graphs. Each of the analyzed groups are accordingly outlined. The following genes were selected and are given as Figures 253 to 271:

GeneID	gene symbol	feature
201105_at	LGALS1	ALL t(4;11) high
204044_at	QPRT	ALL t(4;11) high
205899_at	CCNA1	ALL t(4;11) high
209168_at	GPM6B	ALL t(4;11) high
213539_at	CD3D	T-ALL high
213894_at	KIAA0960	ALL t(4;11) high
215925_s_at		ALL t(4;11) high
218224_at	PNMA1	T-ALL high
219463_at	C20orf103	ALL t(4;11) high
219631_at	FLJ12929	T-ALL high
225563_at	ESTs	ALL t(4;11) high

225592_at	NRM	ALL t(4;11) high
228083_at	Homo sapiens mRNA; cDNA DKFZp434I1216 (from clone DKFZp434I1216)	ALL t(4;11) high
228988_at	ZNF6	T-ALL high
235749_at		ALL t(8;14) high
242414_at	ESTs	ALL t(4;11) high
243756_at	ESTs	ALL t(4;11) high

Example 7 - Results VI: Analysis of 8 AML subtypes

Here we analyzed in total 8 distinct AML subtypes. We applied the described two different statistical methods for identification of genes which allow accurate class assignments to the respective groups.

trisomy 8	(n=10)
other aberrant	(n=5)
complex	(n=36)
normal	(n=62)
t(8;21)	(n=13)
t(15;17)	(n=20)
inv(16)	(n=12)
MLL	(n=15)

First, expression data were analyzed according to example 3 as described hereinabove.

A set of 20 top-ranked genes, which provided both optimal classification accuracy and highest prediction strength for all pairwise (all pairs) and one-versus-all comparisons is given as table 35. Within this set of genes, optimal classification accuracy can be obtained with genes marked by asterisks. Gene expression intensities, plotted as bar graphs are given in Figures 272 to 336. Genes are depicted as unique Affymetrix identifier (for example 201497_x_at) and, where available, approved HGNC symbols (HUGO Gene Nomenclature Committee). More detailed, the complete annotation and sequence information about this set of genes is listed in table 43a,b.

In total 173 cases of AML were analyzed. 160 of 174 cases (92.5%) were assigned to the correct AML subtype in all pairwise comparisons (table 34a). The sensitivity indicated for each subgroup indicates the percentage of cases of the specific subgroup identified correctly in all pairwise comparisons (range 60% to 100%). The specificity indicates for each subgroup the percentage of correct assignments to this subgroup (range 85.5% to 100%).

In total 1211 individual assignments of AML were analyzed. 1198 of 1211 assignments (98.9%) were correct (table 34b). The sensitivity indicated for each subgroup indicates the percentage of correct assignments for cases of the specific subgroup in pairwise comparisons (range 94.3% to 100%). The specificity indicates for each subgroup the percentage of correct assignments to this subgroup (range 97.7% to 100%).

In a second approach significant genes were identified according to Westfall & Young. Table 36 represents all genes found to be significant after p-value adjustment. Genes are depicted as unique Affymetrix identifier (for example 201497_x_at) and, where available, approved HGNC symbols (HUGO Gene Nomenclature Committee). More detailed, the complete annotation and sequence information about this set of genes is listed in table 43a,b.

Furthermore, we provide information about genes which were found to be rated significant independently by both methodologies (Table 36). Top-significant genes according to the method of example 3 are marked by asterisks. Genes which were included in any of the top-20 lists are marked by positive signs.

In addition, selected gene profiles were chosen to demonstrate their capability of discriminating different leukemia types, subtypes and normal bone marrow, respectively. Gene expression profiles were generated by means of PERL-programs, evaluated and plotted as bar graphs. Each of the analyzed groups are accordingly outlined. The following genes were selected and are given as Figures 337 to 370:

GeneID	gene symbol	feature
201497_x_at	MYH11	AML inv(16) high
228827_at	Homo sapiens clone 25023 mRNA sequence	AML t(8;21) high
38487_at	FLJ12442	AML t(15;17) high
203074_at	ANXA8	AML t(15;17) high
205528_s_at	CBFA2T1	AML t(8;21) high
205529_s_at	CBFA2T1	AML t(8;21) high
206940_s_at	POU4F1	AML t(8;21) high
211341_at	POU4F1	AML t(8;21) high
201496_x_at	MYH11	AML inv(16) high
228660_x_at	SEMA4F	other high
202718_at	IGFBP2	AML t(15;17) high

205380_at	PDZK1	other high
202746_at		AML MLL low
201596_x_at	KRT18	AML t(8;21) low
34210_at	CDW52	AML t(15;17) low
212850_s_at	LRP4	AML inv(16) high
228904_at	ESTs	AML t(8;21) low, AML t(15;17) low, AML inv(16) low, AML MLL low
203151_at	MAP1A	AML t(8;21) low
201137_s_at	HLA-DPB1	AML t(15;17) low
200675_at	CD81	AML inv(16) low
201425_at	ALDH2	AML t(8;21) low
202085_at	TJP2	AML inv(16) low
202619_s_at	PLOD2	AML MLL low
203092_at	TIMM44	AML inv(16) low
204425_at	ARHGAP4	AML t(15;17) low
205366_s_at	HOXB6	AML t(8;21) low, AML t(15;17) low, AML inv(16) low, AML MLL low
205472_s_at	DACH	AML MLL high
206761_at	TACTILE	AML MLL low

222166_at		AML +8 low
222335_at	ESTs	AML MLL low
223318_s_at	MGC10974	AML complex low
225330_at	Homo sapiens, clone MGC:18216 IMAGE:4156235, mRNA, complete cds	AML inv(16) low
231277_x_at	ESTs	AML complex low
635_s_at	PPP2R5B	other low

Example 7 - Results VII: Analysis of 5 genetically defined CLL subtypes

Here we analyzed in total 5 genetically defined CLL subtypes. We applied the described two different statistical methods for identification of genes which allow accurate class assignments to the respective groups.

trisomy 12 (n=5)

11q- (n=4)

13q- (n=10)

17p- (n=4)

normal (n=9)

First, expression data were analyzed according to example 3 as described hereinabove.

A set of 20 top-ranked genes, which provided both optimal classification accuracy and highest prediction strength for all pairwise (all pairs) and one-versus-all comparisons is given as table 38. Within this set of genes, optimal classification accuracy can be obtained with genes marked by asterisks. Gene expression

intensities, plotted as bar graphs are given in Figures 371 to 404. Genes are depicted as unique Affymetrix identifier (for example 201497_x_at) and, where available, approved HGNC symbols (HUGO Gene Nomenclature Committee). More detailed, the complete annotation and sequence information about this set of genes is listed in table 43a,b.

In total 32 cases of CLL were analyzed. 31 of 32 cases (96.9%) were assigned to the correct CLL subtype in all pairwise comparisons (table 37a). The sensitivity indicated for each subgroup indicates the percentage of cases of the specific subgroup identified correctly in all pairwise comparisons (range 90% to 100%). The specificity indicates for each subgroup the percentage of correct assignments to this subgroup (range 90% to 100%).

In total 128 individual assignments of CLL were analyzed. 127 of 128 assignments (99.2%) were correct (table 37b). The sensitivity indicated for each subgroup indicates the percentage of correct assignments for cases of the specific subgroup in pairwise comparisons (range 97.5% to 100%). The specificity indicates for each subgroup the percentage of correct assignments to this subgroup (range 97.3% to 100%).

In a second approach significant genes were identified according to Westfall & Young. Table 39 represents all genes found to be significant after p-value adjustment. Genes are depicted as unique Affymetrix identifier (for example 201497_x_at) and, where available, approved HGNC symbols (HUGO Gene Nomenclature Committee). More detailed, the complete annotation and sequence information about this set of genes is listed in table 43a,b.

Furthermore, we provide information about genes which were found to be rated significant independently by both methodologies (Table 39). Top-significant genes according to the method of example 3 are marked by asterisks. Genes which were included in any of the top-20 lists are marked by positive signs.

Example 7 - Results VIII: Analysis of the four major leukemia types (ALL, AML, CLL, CML) and normal bone marrow

Here we analyzed in total 4 major leukemia types as well a cohort of healthy volunteers for normal bone marrow characteristics. We applied the described two different statistical methods for identification of genes which allow accurate class assignments to the respective groups.

ALL (n=47)

AML (n=175)

CLL (n=35)

CML (n=14)

Normal bone marrow (n=9)

First, expression data were analyzed according to example 3 as described hereinabove.

A set of 20 top-ranked genes, which provided both optimal classification accuracy and highest prediction strength for all pairwise (all pairs) and one-versus-all comparisons is given as table 41. Within this set of genes, optimal classification accuracy can be obtained with genes marked by asterisks. Gene expression intensities, plotted as bar graphs are given in Figures 405 to 431. Genes are depicted as unique Affymetrix identifier (for example 201497_x_at) and, where available, approved HGNC symbols (HUGO Gene Nomenclature Committee). More detailed, the complete annotation and sequence information about this set of genes is listed in table 43a,b.

In total 280 cases of leukemia and normal bone marrow (BM) were analyzed. 263 of 280 cases (93.9%) were assigned to the correct leukemia subtype or normal bone marrow in all pairwise comparisons (table 40a). The sensitivity indicated for each subgroup indicates the percentage of cases of the specific subgroup identified correctly in all pairwise comparisons (range 76.6% to 98.3%). The

specificity indicates for each subgroup the percentage of correct assignments to this subgroup (range 88.9% to 97.1%).

In total 1120 individual assignments of leukemia subtype or normal bone marrow were analyzed. 1103 of 1120 assignments (98.5%) were correct (table 40b). The sensitivity indicated for each subgroup indicates the percentage of correct assignments for cases of the specific subgroup in pairwise comparisons (range 94.2% to 99.3%). The specificity indicates for each subgroup the percentage of correct assignments to this subgroup (range 97.2% to 99.3%).

In a second approach significant genes were identified according to Westfall & Young. Table 42 represents all genes found to be significant after p-value adjustment. Genes are depicted as unique Affymetrix identifier (for example 201497_x_at) and, where available, approved HGNC symbols (HUGO Gene Nomenclature Committee). More detailed, the complete annotation and sequence information about this set of genes is listed in table 43a,b.

Furthermore, we provide information about genes which were found to be rated significant independently by both methodologies (Table 42). Top-significant genes according to the method of example 3 are marked by asterisks. Genes which were included in any of the top-20 lists are marked by positive signs.

In addition, selected gene profiles were chosen to demonstrate their capability of discriminating different leukemia types, subtypes and normal bone marrow, respectively. Gene expression profiles were generated by means of PERL-programs, evaluated and plotted as bar graphs. Each of the analyzed groups are accordingly outlined. The following genes were selected and are given as Figures 432 to 464:

GeneID	gene symbol	feature
202503_s_at	KIAA0101	CLL low

202580_x_at	FOXM1		CLL low
202709_at	FMOD		CLL high
204882_at	KIAA0053		CLL high
205049_s_at	CD79A		ALL high, CLL high
205051_s_at	KIT		AML high
205382_s_at	DF		AML high
205599_at	TRAF1		CML low CLL high
206255_at	BLK		ALL high, CLL high
206398_s_at	CD19		ALL high, CLL high
210487_at	DNTT		ALL high
210948_s_at	LEF1		ALL high, CLL high
211352_s_at	NCOA3		CLL high
211404_s_at	APLP2		AML high
214761_at	OAZ		ALL high
217950_at	NOSIP		CLL high
218090_s_at			CLL high
218516_s_at	FLJ20421		normal BM low
218916_at	FLJ23436		normal BM low

219753_at	STAG3	ALL high
221969_at	PAX5	ALL high, CLL high
223703_at	CDA017	AML high, CML high, normal BM high
226147_s_at	Homo sapiens cDNA: FLJ22667 fis, clone HSI08385	CLL high
228471_at	ESTs	CLL high
229487_at	ESTs	ALL high
229790_at	TERF2	CML low, BM low
231736_x_at	MGST1	AML high, CML high, normal BM high
231854_at	Homo sapiens cDNA FLJ11448 fis, clone HEMBA1001391	CML low
239287_at	ESTs	CLL high
243362_s_at	LEF1	ALL high
243363_at	LEF1	ALL high, CLL high
41577_at	PPP1R16B	CML low

Tables 43a, b: functional gene annotation for genes identified to be differentially expressed between different types of leukemia, or between healthy bone marrow and leukemia, respectively.

As described by the GeneChip manufacturer, for each probeset (for example 200093_s_at_HG-U133A), a GenBank or RefSeq accession number was chosen

to represent the target sequence. Using this accession number, a UniGene cluster (in current release) was identified where the accession number was used. If there is a link to LocusLink in the UniGene record, then annotations were retrieved from LocusLink. Those annotations include gene symbol, location, OMIM, EC, Gene Ontology (GO), description and RefSeq sequence accession. The RefSeq accession was linked to the protein annotations, which include domain identification (Pfam and BLOCKS), similarity search (blastp nr) and family classification (SCOP, EC and GPCR HMM searches).

Target sequence information for all the probes which were identified to be able to distinguish between different types and subtypes of leukemia and normal bone marrow, respectively, are given in Table 44.

As further described by the GeneChip manufacturer, the HG-U133 Target Databank is a compilation of probe set annotations and target sequence information for all the probes represented on the HG-U133 A and B arrays. Target sequences are the relatively short (typically around 300-600 bp) sequences against which probes have been designed on a GeneChip® array. These target sequences can be thought of as a subsequence of the Consensus/Exemplar sequence.

The Consensus/Exemplar sequences (i.e., the coding or full cDNA sequences corresponding to the markers described herein as being able to distinguish between different types and subtypes of leukemia and normal bone marrow) for most markers are given in Table 45.

Example 7 Conclusions

The expression pattern of genes allowed precise class assignments of defined leukemia types and subtypes according to the WHO classification of hematological malignancies, and normal BM, respectively.

Thus, we introduce candidate genes suitable for diagnosis of leukemia types and subtypes based on gene expression profiling.

These data demonstrate the utility of gene expression profiling for the discrimination of all leukemia major entities and most subentities. In total, up to 14 different leukemia types and subtypes could clearly be distinguished from each other and from normal BM, respectively. These leukemias are associated with highly differing prognoses and require specific treatment strategies. By performing these analyses on a single platform requiring basic molecular biological methods, this approach provides a broad access to high-quality diagnosis of leukemia.

Golub			
A -	samples: 18 / 85		
accuracy	0,87		
confidence	0,77		
failed	6,19,22,26,78,79,80,81,82,83,84,85,99		
gene	signal-to-noise	p	decision limit
g1	-1,14	0*	482,01
g2	-1,06	0*	192,17
g3	-0,97	0*	207,67
g4	0,94	0*	205,05
g5	-0,93	0*	1818,11
g6	0,93	0*	451,74
g7	-0,91	0*	23,84
g8	-0,90	0*	225,72
g9	0,90	0*	43,85
g10	0,89	0*	210,78
g11	-0,88	0*	118,63
g12	0,87	0*	55,39

invention			
A -	samples: 18 / 85		
accuracy	0,96		
confidence	0,88		
failed	5,6,19,22		
gene	signal-to-noise	p	decision limit
g1	-1,14	0	
g2	-1,06	0*	98,50
g3	-0,97	0	
g4	0,94	0	
g5	-0,93	0	
g6	0,93	0	
g7	-0,91	0	
g8	-0,90	0	
g9	0,90	0	
g10	0,89	0	
g11	-0,88	0	
g12	0,87	0*	67,80

g13	0,87	0*	127,15	g13	0,87	0*	164,10
g14	0,86	0*	222,04	g14	0,86	0	
g15	0,85	0*	68,52	g15	0,85	0	
g16	-0,85	0*	546,97	g16	-0,85	0	
g17	0,84	0*	1242,77	g17	0,84	0	
g18	-0,84	0*	162,61	g18	-0,84	0	
g19	-0,83	0*	385,39	g19	-0,83	0	
g20	0,46	0*	105,38	g20	0,46	0	

Table A. Analysis of 18 samples class A versus 85 samples class non-A. On the left the analysis according to Golub is presented for 20 informative genes. The crossvalidation accuracy is 0,87, confidence 0,77. Samples, where crossvalidation failed, are listed. For each gene signal to noise ratio, p-value (significance obtained from permutation test) and decision limit are provided. On the right the same data set is analyzed using the protocol of the invention. By selection of 3 genes (marked with asterisks) out of the top 20 genes and selecting optimized decision limits, the crossvalidation accuracy reaches 0,96, confidence 0,88.

Table 1

UCL/HGNC/HUGO Human Gene Nomenclature Database Symbol	GenBank Accession No.	UniGene Cluster	RefSeq	Chromosomal Location	Description Unigene Build #95	34210_at
GAS2-related on chromosome 22	Y07846	Hs.322852	NM_006478	22q11.2	Cluster Incl. Y07846:H.sapiens mRNA for GAR22 protein /cds=(132,1145) /gb=Y07846 /gi=1666070 /ug=Hs.15346 /len=2238	31874_at
HLA-DPA1	X00457	Hs.914		6p21.3	Cluster Incl. X00457:Human mRNA for SB classII histocompatibility antigen alpha-chain /cds=(0,702) /gb=X00457 /gi=36405 /ug=Hs.914 /len=1048	38833_at
ADRA2C (adrenergic, alpha-2C-, receptor)	J03853	Hs.123022	NM_000683	4p16	Cluster Incl. J03853:Human kidney alpha-2-adrenergic receptor mRNA, complete cds /cds=(38,1423) /gb=J03853 /gi=178193 /ug=Hs.123022 /len=1491	34512_at

POU4F1 (POU domain, class 4, transcription factor 1)	X64624	Hs.211588	NM_006237	13q21.1-q22	Cluster Incl. X64624:H.sapiens mRNA for RDC-1 POU domain containing protein /cds=(277,1272) /gb=X64624 /gi=35914 /ug=Hs.211588 /len=3492	35940_at
CLECSF2 (C-type (calcium dependent, carbohydrate-recognition domain) lectin, superfamily member 2 (activation-induced))	X96719	Hs.85201	NM_005127	12p13-p12	Cluster Incl. X96719:H.sapiens mRNA for AICL (activation-induced C-type lectin) /cds=(132,581) /gb=X96719 /gi=1632815 /ug=Hs.85201 /len=739	40698_at
PTGDS (prostaglandin D2 synthase (21kD, brain))	AI207842	Hs.8272	NM_000954	9q34.2-q34.3	Cluster Incl. AI207842:ao89h09.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-1953089 /clone_end=3 /gb=AI207842 /gi=3769784 /ug=Hs.8272 /len=771	38407_r_at
TRH (thyrotropin-releasing hormone)	M63582	Hs.182231	NM_007117	3q13.3-q21	Cluster Incl. M63582:Human preprothyrotropin-releasing hormone gene /cds=(8,736) /gb=M63582 /gi=190297 /ug=Hs.182231 /len=1457	32323_at
DKFZP586N1922	N99340	Hs.7357		19	Cluster Incl. N99340:IMAGE-20074 Homo sapiens cDNA /clone=IMAGE-20074	36095_at

					/gb=N99340 /gi=1270755 /ug=Hs.7357 /len=1110	
PTGDS (prostaglandin D2 synthase	M98539	Hs.8272	NM_000954	9q34.2-q34.3	M98539 /FEATURE=exon /DEFINITION=HUMPDS03 Human prostaglandin D2 synthase gene, exon 7	216_at
HLA-DQB1	M81141	Hs.73931	NM_002123	6p21.3	Cluster Incl. M81141:Human MHC class II HLA-DQ-beta mRNA (DR7 DQw2), complete cds /cds=(35,820) /gb=M81141 /gi=188202 /ug=Hs.73933 /len=1171	36773_f_at
CTSW (cathepsin W (lymphopain))	AF013611	Hs.87450	NM_001335	11q13.1	Cluster Incl. AF013611:Homo sapiens lymphopain mRNA, complete cds /cds=(0,1130) /gb=AF013611 /gi=2582044 /ug=Hs.87450 /len=1131	40718_at
KIAA0246	D87433	Hs.301989	NM_015136	3	Cluster Incl. D87433:Human mRNA for KIAA0246 gene, partial cds /cds=(0,6639) /gb=D87433 /gi=1665760 /ug=Hs.84753 /len=6777	38487_at

MYH11	AF013570	Hs.78344	NM_002474, NM_022844, NM_022870	16p13.13-p13.12	Cluster Incl. AF013570:Homo sapiens smooth muscle myosin heavy chain SM2 mRNA, alternatively spliced, partial cds /cds=(0,1767) /gb=AF013570 /gi=2352944 /ug=Hs.78344 /len=2580	37407_s_at
KRT18 (keratin 18)	M26326	Hs.65114	NM_000224	12q13	Cluster Incl. M26326:Human keratin 18 mRNA, complete cds /cds=(51,1343) /gb=M26326 /gi=186690 /ug=Hs.65114 /len=1412	35766_at
POU4F1 (POU domain, class 4, transcription factor 1)	L20433	Hs.211588	NM_006237	13q21.1-q22	Cluster Incl. L20433:Human octamer binding transcription factor 1 (OTF1) mRNA, complete cds /cds=(234,1496) /gb=L20433 /gi=418015 /ug=Hs.211588 /len=3824	35939_s_at
CRA	U78556	Hs.166066	NM_006697	1	U78556 /FEATURE= /DEFINITION=HSU78556 Human cisplatin resistance associated alpha protein (hCRA alpha) mRNA, complete cds	1230_g_at

MYH11	AF001548	Hs.78344	NM_002474	16p13.13-p13.12	AF001548 /FEATURE=mRNA /DEFINITION=HUAF001548 Human Chromosome 16 BAC clone CIT987SK-A-815A9, complete sequence	767_at

Table 2:

UCL/HGNC/HUGO Human Gene Nomenclature Database Symbol	GenBank Accession No.	UniGene Cluster	RefSeq	Chromosomal Location	Description Unigene Build #95	
HOXB2 (homeo box B2)	X16665	Hs.2733	NM_002145	17q21-q22	Cluster Incl. X16665:Human HOX2H mRNA from the Hox2 locus /cds=(78,1148) /gb=X16665 /gi=32381 /ug=Hs.2733 /len=1520	39610_at
CTSW (cathepsin W (lymphopain))	AF013611	Hs.87450	NM_001335	11q13.1	Cluster Incl. AF013611:Homo sapiens lymphopain mRNA, complete cds /cds=(0,1130) /gb=AF013611 /gi=2582044 /ug=Hs.87450 /len=1131	40718_at
C-type (calcium dependent, carbohydrate-recognition domain) lectin, superfamily member 2 (activation-induced)/ug=Hs.85201 /len=739	X96719	Hs.85201	NM_005127	12p13-p12	Cluster Incl. X96719:H.sapiens mRNA for AICL (activation-induced C-type lectin) /cds=(132,581) /gb=X96719 /gi=1632815 /ug=Hs.85201 /len=739	40698_at

myosin, heavy polypeptide 11, smooth muscle	AF001548	Hs.78344	NM_002474	16p13.13-p13.12	AF001548 /FEATURE=mRNA /DEFINITION=HUAF001548 Human Chromosome 16 BAC clone CIT987SK-A-815A9, complete sequence	767_at
Human mRNA for SB classII histocompatibility antigen alpha-chain	X00457	Hs.914	NM_033554	6p21.3	Cluster Incl. X00457:Human mRNA for SB classII histocompatibility antigen alpha-chain /cds=(0,702) /gb=X00457 /gi=36405/ug=Hs.914 /len=1048	38833_at
	AF013570	Hs.78344	NM_002474	16p13.13-p13.12	Cluster Incl. AF013570:Homo sapiens smooth muscle myosin heavy chain SM2 mRNA, alternatively spliced, partial cds /cds=(0,1767) /gb=AF013570 /gi=2352944 /ug=Hs.78344 /len=2580	37407_s_at
ARHGAP4 (Rho GTPase activating protein 4)	X78817	Hs.3109	NM_001666	xq28	Cluster Incl. X78817:H.sapiens partial C1 mRNA /cds=(42,2882) /gb=X78817 /gi=840785 /ug=Hs.3109 /len=3236	39649_at
Homo sapiens mRNA for KIAA0246 protein, partial	D87433	Hs.84753			Cluster Incl. D87433:Human mRNA for KIAA0246 gene, partial cds /cds=(0,6639)	38487_at

cdsgi|1665760|dbj|D87433.1|D87433[1665760]					/gb=D87433 /gi=1665760 /ug=Hs.84753/len=6777	
major histocompatibility complex, class II, DM alpha	X62744	Hs.77522	NM_006120	6p21.3	Cluster Incl. X62744:Human RING6 mRNA for HLA class II alpha chain-like product /cds=(45,830) /gb=X62744 /gi=36062 /ug=Hs.77522 /len=1079	37344_at
voltage-dependent anion channel 1	L06132	Hs.149155	NM_003374	Xq13-q21	Cluster Incl. L06132:Human voltage-dependent anion channel isoform 1 (VDAC) mRNA, complete cds /cds=(99,950) /gb=L06132 /gi=340198 /ug=Hs.149155 /len=1806	40198_at
ITGB2 (integrin, beta 2 (antigen CD18 (p95), lymphocyte function-associated antigen ; macrophageantigen 1 (mac-1) beta subunit))	M15395	Hs.83968	NM_000211	21q22.3	Cluster Incl. M15395:Human leukocyte adhesion protein (LFA-1/Mac-1/p150,95 family) beta subunit mRNA /cds=(72,2381 /gb=M15395 /gi=186933 /ug=Hs.83968 /len=2776	37918_at
SERPING1 (serine (or cysteine) proteinase inhibitor, clade G (C1 inhibitor), member 1,	X54486	Hs.151242	NM_000062	11q12-q13.1	Cluster Incl. X54486:Human gene for C1-inhibitor /cds=(60,1562) /gb=X54486	39775_at

(angioedema, hereditary))					/gi=29534 /ug=Hs.151242 /len=1827	
Homo sapiens mRNA; cDNA DKFZp564K0822 (from clone DKFZp564K0822)	W25986	Hs.4750		6	Cluster Incl. W25986:17e7 Homo sapiens cDNA /gb=W25986 /gi=1306253 /ug=Hs.4750 /len=769	34830_at
CBFA2T1 (core-binding factor, runt domain, alpha subunit 2; translocated to, 1; cyclin D-related)	D43638	Hs.31551	NM_004349	8q22	Cluster Incl. D43638:Human mRNA for MTG8a protein, complete cds /cds=(411,2144) /gb=D43638 /gi=940399 /ug=Hs.31551 /len=3460	35638_at
DKFZP586N1922 protein	N99340	Hs.7357		19	Cluster Incl. N99340:IMAGE-20074 Homo sapiens cDNA /clone=IMAGE-20074 /gb=N99340 /gi=1270755 /ug=Hs.7357 /len=1110	36095_at
ADP-ribosylation factor related protein	X91504	Hs.64904	NM_003224	20q13	Cluster Incl. X91504:H.sapiens mRNA for ARP1 protein/cds=(11,616) /gb=X91504 /gi=1103581 /ug=Hs.64904 /len=1541	35765_at

HLA-DPB1 (major histocompatibility complex, class II, DP beta 1)	M83664	Hs.814	NM_002121	6p21.3	Cluster Incl. M83664:Human MHC class II lymphocyte antigen (HLA-DP) beta chain mRNA, complete cds /cds=(59,835) /gb=M83664 /gi=188478 /ug=Hs.814 /len=1501	38096_f_at
ADP-RIBOSYLATION FACTOR-RELATED PROTEIN 1; ARFRP1	X91504	Hs.64904	NM_003224	20q13.3	Cluster Incl. X91504:H.sapiens mRNA for ARP1 protein/cds=(11,616) /gb=X91504 /gi=1103581 /ug=Hs.64904 /len=1541	36142_at
HLA-DPB1 (major histocompatibility complex, class II, DP beta 1)	M83664	Hs.814	NM_002121	6p21.3	Cluster Incl. M83664:Human MHC class II lymphocyte antigen (HLA-DP) beta chain mRNA, complete cds /cds=(59,835) /gb=M83664 /gi=188478 /ug=Hs.814 /len=1501	38095_i_at
annexin V	U05770	Hs.79274		4q26-q28	Cluster Incl. U05770:Human annexin V (ANX5) gene /cds=(164,1126/gb=U05770 /gi=2182176 /ug=Hs.79274 /len=1597	37747_at
CD74 (CD74 antigen (invariant polypeptide of major histocompatibility complex, class II	M13560	Hs.84298	NM_004355	5q32	Cluster Incl. M13560:Human Ia-associated invariant gamma-chain gene	35016_at

antigen-associated))					/cds=(795,1493) /gb=M13560 /gi=184518 /ug=Hs.84298 /len=2080	
interleukin 13 receptor, alpha 1	Y10659	Hs.285115	NM_001560	X	Y10659 /FEATURE=cds /DEFINITION=HSIL13RA. H.sapiens IL-13Ra mRNA	359_at
meningioma (disrupted in balanced translocation) 1	X82209	Hs.79085	NM_002430	22q12.1	Cluster Incl. X82209:H.sapiens MN1 mRNA /cds=(887,4915)/gb=X82209 /gi=804991 /ug=Hs.79085 /len=7554	37283_at
CDw52, cell surface	N90866	Hs.276770	NM_001803	1p36	Cluster Incl. N90866:zb11b10.s1 Homo sapiens cDNA, 3 end /clone=IMAGE-301723 /clone_end=3 /gb=N90866 /gi=1444193 /ug=Hs.214742 /len=577	34210_at
transforming growth factor, beta-induced, 68kD	M77349	Hs.118787	NM_000358	5q31	M77349 /FEATURE= /DEFINITION=HUMTGFBIG Human transforming growth factor-beta induced gene product (BIGH3) mRNAcomplete cd	1385_at

MGC2747(hypothetical protein MGC2747	AL046940	Hs.250723	NM_024104	19	Cluster Incl. AL046940:DKFZp586I0517_r1 Homo sapiens cDNA, 5 end /clone=DKFZp586I0517 /clone_end=5 /gb=AL046940 /gi=5434999 /ug=Hs.231657 /len=695	41273_at
major histocompatibility complex, class II, DR alpha	J00194	Hs.76807	NM_019111	6p21.3	Cluster Incl. J00194:human hla-dr antigen alpha-chainmma & ivs fragments /cds=(26,790) /gb=J00194 /gi=188231/ug=Hs.76807 /len=1199	37039_at
bone gamma-carboxyglutamate (gla) protein (osteocalc	AI131030	Hs.2558	NM_000711	1q25-q31	Cluster Incl. AI131030:qb82f10.x1 Homo sapiens cDNA3' end /clone=IMAGE-1706635 /clone_end=3' /gb=AI131030/gi=3601046 /ug=Hs.2558 /len=565	36253_at
NPR3 (natriuretic peptide receptor C/guanylate cyclase C (atrionatriuretic peptide receptor C))	M59305	Hs.123655	NM_000908	5p14-p13	Cluster Incl. M59305:Human atrial natriuretic peptide clearance receptor (ANP C-receptor) mRNA, complete cds /cds=(362,1987) /gb=M59305 /gi=178651	34519_at

					/ug=Hs.123655 /len=2081	
aldehyde dehydrogenase 2, mitochondrial	X05409	Hs.195432	NM_000690	12q24.2	Cluster Incl. . X05409:Human RNA for mitochondrial aldehyde dehydrogenase I ALDH I (EC 1.2.1.3) /cds=(36,1586) /gb=X0540/gi=28605 /ug=Hs.195432 /len=1989	32747_at

Table 3:

UCL/HGNC/HUGO Human Gene Nomenclature Database Symbol	GenBank Accession No.	UniGene Cluster	RefSeq	Chromosomal Location	Description Unigene Build #95	Gene Name
CAMP (cathelicidin antimicrobial peptide)	Z38026	Hs.51120	NM_004345	3p21:3	Cluster Incl. Z38026:H.sapiens mRNA for FALL-39 peptide antibiotic /cds=(11,523) /gb=Z38026 /gi=558378 /ug=Hs.51120 /len=615	36710_at
SYNE-1B(synaptic nuclear envelope 1)]	AB018339	Hs.8182		6	Cluster Incl. AB018339:Homo sapiens mRNA for KIAA0796 protein, partial cds /cds=(0,3243) /gb=AB018339 /gi=3882312 /ug=Hs.8182 /len=3900	38113_at
ALDH1A1 (aldehyde dehydrogenase 1 family, member A1)	K03000	Hs.76392	NM_000689	9q21	Cluster Incl. K03000:Human aldehyde dehydrogenase 1 mRNA /cds=(0,1022) /gb=K03000 /gi=178399 /ug=Hs.76392 /len=1560	37015_at

LIG1 (ligase I, DNA, ATP-dependent)	AL039458	Hs.4193		3p14	Cluster Incl. AL039458:DKFZp434N0910_s1 Homo sapiens cDNA, 3 end /clone=DKFZp434N0910 /clone_end=3 /gb=AL039458 /gi=5408506 /ug=Hs.4193 /len=849	34800_at
SGP28(specific granule protein (28 kDa); cysteine-rich secretory protein-3)	X94323	Hs.54431	NM_006061	6	Cluster Incl. X94323:H.sapiens mRNA for SGP28 protein /cds=(40,777) /gb=X94323 /gi=1213612 /ug=Hs.54431 /len=2124	36464_at
CBX7 (chromobox homolog 7)	AL031846			22q13.1	Cluster Incl. AL031846:dJ742C19.5 (novel Chromobox protein) /cds=(89,844) /gb=AL031846 /gi=4164368 /ug=Hs.7442 /len=3964	36894_at
PPBP (pro-platelet basic protein (includes platelet basic protein, beta-thromboglobulin, connective	M54995	Hs.2164	NM_002704	4q12-q13	Cluster Incl. M54995:Human connective tissue activation peptide III mRNA, complete cds /cds=(66,452) /gb=M54995 /gi=181175 /ug=Hs.2164 /len=673	39209_r_at

KCNH2 (potassium voltage-gated channel, subfamily H (eag-related), member 2)	AF052728	Hs.188021	NM_000238	7q35-q36	Cluster Incl. AF052728:Homo sapiens HERG-USO (HERG) mRNA, alternatively spliced, partial cds /cds=(0,284) /gb=AF052728 /gi=3549258 /ug=Hs.165664 /len=767	38225_at
PPBP (pro-platelet basic protein (includes platelet basic protein, beta-thromboglobulin, connective	M54995	Hs.2164	NM_002704	4q12-q13	Cluster Incl. M54995:Human connective tissue activation peptide III mRNA, complete cds /cds=(66,452) /gb=M54995 /gi=181175 /ug=Hs.2164 /len=673	39208_i_at
PF4 (platelet factor 4)	M25897	Hs.81564	NM_002619	4q12-q21	M25897 /FEATURE=mRNA /DEFINITION=HUMPF4A Human platelet factor 4 (PF4) mRNA, complete cds	1115_at
PLSCR1 (phospholipid scramblase 1)	AB006746	Hs.198282	NM_021105	3q23	Cluster Incl. AB006746:Homo sapiens hMmTRA1b mRNA, complete cds /cds=(256,1212) /gb=AB006746 /gi=3510296 /ug=Hs.198282 /len=2077	32775_r_at
LCN2 (lipocalin 2 (oncogene 24p3))	AI762213	Hs.204238	NM_005564	9q34	Cluster Incl. AI762213:wi54d04.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-	32821_at

					2394055 /clone_end=3 /gb=AI762213 /gi=5177880 /ug=Hs.204238 /len=677	
PLCE2 (phospholipase C, epsilon 2)	AB029015	Hs.54886		3p25.3-p25.1	Cluster Incl. AB029015:Homo sapiens mRNA for KIAA1092 protein, partial cds /cds=(0,3464) /gb=AB029015 /gi=5689520 /ug=Hs.54886 /len=4147	41796_at
MMP9 (matrix metalloproteinase 9 (gelatinase B, 92kD gelatinase, 92kD type IV collagenase))	J05070	Hs.151738	NM_004994	20q11.2-q13.1	Cluster Incl. J05070:Human type IV collagenase mRNA, complete cds /cds=(19,2142) /gb=J05070 /gi=177204 /ug=Hs.151738 /len=2334	31859_at
TRB@ (T cell receptor beta locus)	M12886	Hs.303157		7q35	M12886 /FEATURE= /DEFINITION=HUMTCBYY Human T-cell receptor active beta-chain mRNA, complete cds	1105_s_at
SPTA1 (spectrin, alpha, erythrocytic 1 (elliptocytosis 2)	M61877	Hs.1985	NM_003126	1q21	Cluster Incl. M61877:Human erythroid alpha-spectrin (SPTA1) mRNA, complete cds /cds=(186,7475) /gb=M61877	38906_at

					/gi=338437 /ug=Hs.1985 /len=8001	
SYNE-2(synaptic nuclei expressed gene 2)	AL080133	Hs.57749	NM_015180	14	Cluster Incl. AL080133:Homo sapiens mRNA; cDNA DKFZp434G173 (from clone DKFZp434G173) /cds=(122,3400) /gb=AL080133 /gi=5262573 /ug=Hs.57749 /len=4307	41815_at
HLA-F (major histocompatibility complex, class I, F)	AL022723	Hs.110309	NM_018950	6p21.3	Cluster Incl. AL022723:dJ377H14.9 (major histocompatibility complex, class I, F (CDA12)) /cds=(97,1185) /gb=AL022723 /gi=5002624 /ug=Hs.110309 /len=1303	37420_i_at
SLU7(step II splicing factor SLU7)	AI660656	Hs.76325	NM_006425	5	Cluster Incl. AI660656:wf23c07.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-2351436 /clone_end=3 /gb=AI660656 /gi=4764239 /ug=Hs.76325 /len=522	37006_at
CTSE (cathepsin E)	J05036	Hs.1355	NM_001910	1q31	J05036 /FEATURE=mRNA /DEFINITION=HUMCTSE Human cathepsin E mRNA, complete cds	271_s_at

CPO (coproporphyrinogen oxidase (coproporphyria, harderoporphyria))	D16611	Hs.89866	NM_000097	3q12	Cluster Incl. D16611:Human mRNA for coproporphyrinogen oxidase, complete cds /cds=(93,1157) /gb=D16611 /gi=469488 /ug=Hs.89866 /len=2333	37999_at
LEF1 (lymphoid enhancer-binding factor 1)	AL049409	Hs.44865	NM_016269	4q23-q25	Cluster Incl. AL049409:Homo sapiens mRNA; cDNA DKFZp586H0919 (from clone DKFZp586H0919) /cds=UNKNOWN /gb=AL049409 /gi=4500194 /ug=Hs.44865 /len=1419	36021_at
TFDP1 (transcription factor Dp-1)	L23959	Hs.79353	NM_007111	13q34	Cluster Incl. L23959:Homo sapiens E2F-related transcription factor (DP-1) mRNA, complete cds /cds=(37,1269) /gb=L23959 /gi=414316 /ug=Hs.79353 /len=1440	37757_at
	S67247				Cluster Incl. S67247:smooth muscle myosin heavy chain isoform SMemb [human, umbilical cord, fetal aorta, mRNA Partial, 971 nt] /cds=(0,681) /gb=S67247 /gi=452986 /ug=Hs.2094 /len=971	32838_at

MMP8 (matrix metalloproteinase 8 (neutrophil collagenase))	J05556	Hs.73862	NM_002424	11q22.3	J05556 /FEATURE=mRNA /DEFINITION=HUMCLGNA Homo sapiens collagenase mRNA, complete cds	681_at
MINPP1 (multiple inositol polyphosphate histidine phosphatase, 1)	AL050356	Hs.95907	NM_004897	10q23	Cluster Incl. AL050356:Homo sapiens mRNA; cDNA DKFZp564L2016 (from clone DKFZp564L2016) /cds=UNKNOWN /gb=AL050356 /gi=4914568 /ug=Hs.95907 /len=2396	38325_at
TCF7 (transcription factor 7 (T-cell specific, HMG-box))	X59871	Hs.169294	NM_003202	5q31.1	Cluster Incl. X59871:Human TCF-1 mRNA for T cell factor 1 (splice form C) /cds=(79,885) /gb=X59871 /gi=36789 /ug=Hs.169294 /len=2910	32649_at
NS1-BP(NS1-binding protein)	AB020657	Hs.197298	NM_006469	1	Cluster Incl. AB020657:Homo sapiens mRNA for KIAA0850 protein, complete cds /cds=(630,2558) /gb=AB020657 /gi=4240188 /ug=Hs.197298 /len=3682	33752_at

CCR2 (chemokine (C-C motif) receptor 2)	U95626	Hs.395	NM_000647	3p21	Cluster Incl. U95626:Homo sapiens ccr2b (ccr2), ccr2a (ccr2), ccr5 (ccr5) and ccr6 (ccr6) genes, complete cds, and lactoferrin (lactoferrin) gene, partial cds /cds=(2,1429) /gb=U95626 /gi=2104517 /ug=Hs.105938 /len=1607	37149_s_at
TRB@ (T cell receptor beta locus)	X00437	Hs.303157		7q35	Cluster Incl. X00437:Human mRNA for T-cell specific protein /cds=(37,975) /gb=X00437 /gi=36748 /ug=Hs.2003 /len=1151	32794_g_at
ADD2 (adducin 2 (beta))	U43959	Hs.247423	NM_001617	2p14-p13	Cluster Incl. U43959:Human beta 4 adducin mRNA, alternatively spliced partial cds /cds=(0,938) /gb=U43959 /gi=1172145 /ug=Hs.4852 /len=1284	36052_at
CD59 (CD59 antigen p18-20 (antigen identified by monoclonal antibodies 16.3A5, EJ16, EJ30, EL32	M84349	Hs.119663	NM_000611	11p13	Cluster Incl. M84349:Human transmembrane protein (CD59) gene /cds=(18,404) /gb=M84349 /gi=180150 /ug=Hs.119663 /len=1840	39351_at

TPM1 (tropomyosin 1 (alpha))	M19267	Hs.77899	NM_000366	15q22.1	Cluster Incl. M19267:Human tropomyosin mRNA, complete cds /cds=(286,1140) /gb=M19267 /gi=339943 /ug=Hs.77899 /len=1633	36791_g_at
TNFRSF7 (tumor necrosis factor receptor *superfamily, member 7)*	M63928	Hs.180841	NM_001242	12p13	Cluster Incl. M63928:Homo sapiens T cell activation antigen (CD27) mRNA, complete cds /cds=(100,882) /gb=M63928 /gi=180084 /ug=Hs.180841 /len=1204	38578_at
TRB@ (T cell receptor beta locus)	X00437	Hs.303157		7q35	Cluster Incl. X00437:Human mRNA for T-cell specific protein /cds=(37,975) /gb=X00437 /gi=36748 /ug=Hs.2003 /len=1151	32793_at
PRKWNK1 (protein kinase, lysine deficient 1)	U00946	Hs.184592	NM_018979	12p13.3	Cluster Incl. U00946:Human clone A9A2BRB5 (CAC)n/(GTG)n repeat-containing mRNA /cds=UNKNOWN /gb=U00946 /gi=405048 /ug=Hs.184592 /len=1971	32185_at

PLXNC1 (*plexin C1*)	AF030339	Hs.286229	NM_005761	12	Cluster Incl. AF030339:Homo sapiens receptor for viral semaphorin protein (VESPR) mRNA, complete cds /cds=(249,4955) /gb=AF030339 /gi=3176761 /ug=Hs.184697 /len=5121	32193_at
TRA@ (T cell receptor alpha locus)	M12959	Hs.74647		14q11.2	M12959 /FEATURE= /DEFINITION=HUMTCAXB Human T-cell receptor active alpha-chain mRNA from JM cell line, complete cds	1106_s_at
CPNE3 (copine III)	AB014536	Hs.14158	NM_003909	8p22-q21.3	Cluster Incl. AB014536:Homo sapiens mRNA for KIAA0636 protein, complete cds /cds=(120,1733) /gb=AB014536 /gi=3327085 /ug=Hs.14158 /len=4737	39706_at
MGEA5 (meningioma expressed antigen 5 (hyaluronidase))	AB014579	Hs.5734	NM_012215	10q24.1-q24.3	Cluster Incl. AB014579:Homo sapiens mRNA for KIAA0679 protein, partial cds /cds=(0,2303) /gb=AB014579 /gi=3327171 /ug=Hs.5734 /len=4303	35317_at

NELL2 (nel (chicken)-like 2)	D83018	Hs.79389	NM_006159	12q13.11-q13.12	Cluster Incl. D83018:Homo sapiens mRNA for nel-related protein 2, complete cds /cds=(96,2546) /gb=D83018 /gi=1827484 /ug=Hs.79389 /len=3198	32598_at
MECP2 (methyl CpG binding protein 2 (Rett syndrome))	AJ132917	Hs.3239	NM_004992	Xq28	Cluster Incl. AJ132917:Homo sapiens mRNA for methyl-CpG-binding protein 2 /cds=(75,1535) /gb=AJ132917 /gi=5419676 /ug=Hs.3239 /len=10091	34355_at
TRA@ (T cell receptor alpha locus)	X02883	Hs.74647		14q11.2	X02883 /FEATURE=cds /DEFINITION=HSTCRAC Human gene for T-cell receptor alpha chain C region	432_s_at
BLVRB (biliverdin reductase B (flavin reductase (NADPH)))	D32143	Hs.76289	NM_000713	19q13.1-q13.2	Cluster Incl. D32143:Human mRNA for biliverdin-IXbeta reductase I /cds=(109,729) /gb=D32143 /gi=699602 /ug=Hs.76289 /len=824	37002_at
PRDX2 (peroxiredoxin 2)	L19185	Hs.146354	NM_005809	13q12	Cluster Incl. L19185:Human natural killer cell enhancing factor (NKEFB) mRNA, complete cds /cds=(124,720) /gb=L19185	39729_at

					/gi=440307 /ug=Hs.146354 /len=980	
AIF1 (allograft inflammatory factor 1)	Y14768	Hs.76364	NM_001623	6p21.3	Cluster Incl. Y14768:Homo sapiens DNA, cosmid clones TN62 and TN82 /cds=(10,744) /gb=Y14768 /gi=3805800 /ug=Hs.890 /len=896	40729_s_at
TPM1 (tropomyosin 1 (alpha))	M19267	Hs.77899	NM_000366	15q22.1	Cluster Incl. M19267:Human tropomyosin mRNA, complete cds /cds=(286,1140) /gb=M19267 /gi=339943 /ug=Hs.77899 /len=1633	36790_at
AMPD3 (adenosine monophosphate deaminase (isoform E))	U29926	Hs.83918	NM_000480	11p15	Cluster Incl. U29926:Human AMP deaminase (AMPD3) gene, promoter 1a region /cds=(453,2777) /gb=U29926 /gi=1002661 /ug=Hs.83918 /len=4018	38463_s_at
	AF035315				Cluster Incl. AF035315:Homo sapiens clone 23664 and 23905 mRNA sequence /cds=UNKNOWN /gb=AF035315 /gi=2661077 /ug=Hs.180737 /len=1331	33267_at

CDC2 (cell division cycle 2, G1 to S and G2 to M)	X05360	Hs.184572	NM_001786	10q21.1	X05360 /FEATURE=cds /DEFINITION=HSCDC2 Human CDC2 gene involved in cell cycle control	1803_at
GCLM (glutamate-cysteine ligase, modifier subunit)	L35546	Hs.89709	NM_002061	1p22.1	Cluster Incl. L35546:Homo sapiens gamma-glutamylcysteine synthetase light subunit mRNA, complete cds /cds=(253,1077) /gb=L35546 /gi=530136 /ug=Hs.89709 /len=1610	33163_r_at
NPAT (nuclear protein, ataxia-telangiectasia locus)	D83243	Hs.89385	NM_002519	11q22-q23	Cluster Incl. D83243:Human NPAT mRNA, complete cds /cds=(66,4349) /gb=D83243 /gi=1304113 /ug=Hs.89385 /len=5900	40732_at
KIAA0471(KIAA0471 gene product)	AB007940				Cluster Incl. AB007940:Homo sapiens mRNA for KIAA0471 protein, complete cds /cds=(412,1524) /gb=AB007940 /gi=3413903 /ug=Hs.107325 /len=6834	34445_at
ITK (IL2-inducible T-cell kinase)	L10717	Hs.211576	NM_005546	5q31-q32	L10717 /FEATURE= /DEFINITION=HUMTKTCS Homo sapiens T cell-specific tyrosine kinase mRNA,	1478_at

					complete cds	
TAL1 (T-cell acute lymphocytic leukemia 1 (NOTE: redefinition of symbol))	M63589	Hs.73828	NM_003189	1p32	M63589 /FEATURE=mRNA#5 /DEFINITION=HUMSCL7 Human stem cell leukemia gene product, exon 6	560_s_at
OLR1 (oxidised low density lipoprotein (lectin-like) receptor 1)	AF079167	Hs.77729	NM_002543	12p13.2-p12.3	Cluster Incl. AF079167:untitled /cds=(61,882) /gb=AF079167 /gi=4050003 /ug=Hs.77729 /len=2468	37233_at
	AL080216				Cluster Incl. AL080216:Homo sapiens mRNA; cDNA DKFZp586K1123 (from clone DKFZp586K1123) /cds=UNKNOWN /gb=AL080216 /gi=5262707 /ug=Hs.26837 /len=2204	35187_at
KIAA0922(KIAA0922 protein)	AB023139	Hs.37892	NM_015196		Cluster Incl. AB023139:Homo sapiens mRNA for KIAA0922 protein, partial cds /cds=(0,2372) /gb=AB023139 /gi=4589475 /ug=Hs.37892 /len=2505	39929_at

GZMK (granzyme K (serine protease, granzyme 3; tryptase II))	U26174	Hs.3066	NM_002104	5q11-q12	Cluster Incl. U26174:Human pre-granzyme 3 mRNA, complete cds /cds=(40,834) /gb=U26174 /gi=829637 /ug=Hs.3066 /len=1040	36280_at
	U23852				Cluster Incl. U23852:Human T-lymphocyte specific protein tyrosine kinase p56lck (lck) abberant mRNA, complete cds /cds=(59,1150) /gb=U23852 /gi=775207 /ug=Hs.1765 /len=2129	33238_at
	L47276				L47276 /FEATURE=UTR#1 /DEFINITION=HUMTOPATR Homo sapiens (cell line HL-60) alpha topoisomerase truncated-form mRNA, 3 UTR	904_s_at
TOSO(regulator of Fas-induced apoptosis)	AF057557	Hs.58831	NM_005449	1	Cluster Incl. AF057557:Homo sapiens anti-Fas-induced apoptosis (TOSO) mRNA, complete cds /cds=(19,1191) /gb=AF057557 /gi=3169292	32967_at

					/ug=Hs.238857 /len=1339	
FCN1 (ficolin (collagen/fibrinogen domain-containing) 1)	S80990	Hs.252136	NM_002003	9q34	Cluster Incl. S80990:ficolin [human, uterus, mRNA, 1736 nt] /cds=(532,1512) /gb=S80990 /gi=1911529 /ug=Hs.169237 /len=1723	36447_at
CD3Z (CD3Z antigen, zeta polypeptide (TiT3 complex))	J04132	Hs.97087	NM_000734	1q22-q23	Cluster Incl. J04132:Human T cell receptor zeta-chain mRNA, complete cds /cds=(74,565) /gb=J04132 /gi=623041 /ug=Hs.97087 /len=1472	37078_at
CEACAM1 (carcinoembryonic antigen-related cell adhesion molecule 1 (biliary glycoprotein))	S71326	Hs.50964	NM_001712	19q13.2	Cluster Incl. S71326:BGPc=biliary glycoprotein adhesion molecule {alternatively spliced} [human, HT29 colon carcinoma cell line, mRNA Partial, 1473 nt] /cds=(0,1394) /gb=S71326 /gi=550030 /ug=Hs.50964 /len=1473	36082_at
DEFA4 (defensin, alpha 4, corticostatin)	AI250799	Hs.2582	NM_001925	8p23	Cluster Incl. AI250799:qi36g07.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-	34546_at

					1858620 /clone_end=3 /gb=AI250799 /gi=3847328 /ug=Hs.2582 /len=542	
GCA (grancalcin, EF-hand calcium-binding protein)	M81637	Hs.79381	NM_012198	2p14-q14.3	Cluster Incl. M81637:Human grancalcin mRNA, complete cds /cds=(119,772) /gb=M81637 /gi=183030 /ug=Hs.79381 /len=1652	37556_at
KIAA0275(KIAA0275 gene product)	D87465	Hs.74583	NM_014767	10	Cluster Incl. D87465:Human mRNA for KIAA0275 gene, complete cds /cds=(316,1590) /gb=D87465 /gi=1665814 /ug=Hs.74583 /len=5316	36155_at
IL2RB (interleukin 2 receptor, beta)	M26062	Hs.75596	NM_000878	22q13.1	M26062 /FEATURE= /DEFINITION=HUMIL2RBC Human interleukin 2 receptor beta chain (p70-75) mRNA, complete cds	1365_at
KIAA0513(KIAA0513 gene product)	AB011085	Hs.301658	NM_014732	16	Cluster Incl. AB011085:Homo sapiens mRNA for KIAA0513 protein, complete cds /cds=(631,1866) /gb=AB011085	38735_at

					/gi=3043549 /ug=Hs.85053 /len=7758	
RHAG (Rhesus blood group-associated glycoprotein)	X64594	Hs.169536	NM_000324	6p21.1-p11	Cluster Incl. X64594:H.sapiens mRNA for 50 kDa erythrocyte plasma membrane glycoprotein /cds=(27,1256) /gb=X64594 /gi=31194 /ug=Hs.169536 /len=1891	32663_at
IGF2R (insulin-like growth factor 2 receptor)	Y00285	Hs.76473	NM_000876	6q26	Y00285 /FEATURE=cds /DEFINITION=HSIGFIIR Human mRNA for insuline-like growth factor II receptor /NOTE=replacement of probe set 972_s_at	160027_s_at
PPP2R5C (protein phosphatase 2, regulatory subunit B (B56), gamma isoform)	U37352	Hs.171734	NM_002719	3p21	U37352 /FEATURE= /DEFINITION=HSU37352 Human protein phosphatase 2A B alpha1 regulatory subunit mRNA, complete cds	176_at
CCR7 (chemokine (C-C motif) receptor 7)	L31584	Hs.1652	NM_001838	17q12-q21.2	L31584 /FEATURE=exon /DEFINITION=HUMEBI103 Human G protein-coupled receptor (EBI 1) gene	1097_s_at

					exon 3, complete cds	
LCK (lymphocyte-specific protein tyrosine kinase)	M36881	Hs.1765	NM_005356	1p35-p34.3	M36881 /FEATURE=mRNA /DEFINITION=HUMLCKAA Human lymphocyte-specific protein tyrosine kinase (lck) mRNA, complete cds	2059_s_at
ALS2CR3 (amyotrophic lateral sclerosis 2 (juvenile) chromosome region, candidate 3)	AB011121	Hs.154248	NM_015049	2q33	Cluster Incl. AB011121:Homo sapiens mRNA for KIAA0549 protein, partial cds /cds=(0,1409) /gb=AB011121 /gi=3043621 /ug=Hs.154248 /len=4745	40064_at
	AF055029				Cluster Incl. AF055029:Homo sapiens clone 24711 mRNA sequence /cds=UNKNOWN /gb=AF055029 /gi=3005759 /ug=Hs.4988 /len=1816	34866_at
GG2-1(TNF-induced protein)	AF099935	Hs.17839	NM_014350	5	Cluster Incl. AF099935:Homo sapiens MDC-3.13 isoform 2 mRNA, complete cds /cds=(84,680) /gb=AF099935 /gi=3860092 /ug=Hs.17839 /len=1897	33243_at

EPOR (erythropoietin receptor)	M60459	Hs.89548	NM_000121	19p13.3-p13.2	Cluster Incl. M60459:Human erythropoietin receptor mRNA, complete cds /cds=(105,1631) /gb=M60459 /gi=182244 /ug=Hs.89548 /len=1818	37986_at
CDC25B (cell division cycle 25B)	S78187	Hs.153752	NM_004358	20p13	S78187 /FEATURE= /DEFINITION=S78187 CDC25Hu2=cdc25+ homolog [human, mRNA, 3118 nt]	1347_at
KLRB1 (killer cell lectin-like receptor subfamily B, member 1)	U11276	Hs.169824	NM_002258	12p13	Cluster Incl. U11276:Human hNKR-P1a protein (NKR-P1A) mRNA, complete cds /cds=(60,737) /gb=U11276 /gi=538270 /ug=Hs.169824 /len=738	35449_at
KIAA0349(KIAA0349 protein)	AB002347			6	Cluster Incl. AB002347:Human mRNA for KIAA0349 gene, partial cds /cds=(0,3827) /gb=AB002347 /gi=2224638 /ug=Hs.15303 /len=6158	39797_at
GYPB (glycophorin B (includes Ss blood group))	U05255	Hs.250653	NM_002100	4q28-q31	Cluster Incl. U05255:Human glycophorin HeP2 mRNA, partial cds /cds=(0,302)	41026_f_at

					/gb=U05255 /gi=454085 /ug=Hs.93223 /len=338	
CREME9(cytokine receptor-like molecule 9)	AF046059	Hs.7120	NM_015986	17	Cluster Incl. AF046059:Homo sapiens cytokine receptor related protein 4 (CYTOR4) mRNA, complete cds /cds=(22,1350) /gb=AF046059 /gi=4105471 /ug=Hs.119410 /len=2848	37509_at
DGKA (diacylglycerol kinase, alpha (80kD))	X62535	Hs.172690	NM_001345	12q13.3	Cluster Incl. X62535:H.sapiens mRNA for diacylglycerol kinase /cds=(103,2310) /gb=X62535 /gi=30822 /ug=Hs.172690 /len=2564	32716_at
KIAA0008(KIAA0008 gene product)	D13633	Hs.77695	NM_014750	14	Cluster Incl. D13633:Human mRNA for KIAA0008 gene, complete cds /cds=(121,2418) /gb=D13633 /gi=286012 /ug=Hs.77695 /len=2640	37231_at
MAF (v-maf musculoaponeurotic fibrosarcoma (avian) oncogene homolog)	AF055376	Hs.30250	NM_005360	16q22-q23	Cluster Incl. AF055376:Homo sapiens short form transcription factor C-MAF (c-maf) mRNA, complete cds	41504_s_at

					/cds=(807,1928) /gb=AF055376 /gi=3335147 /ug=Hs.30250 /len=4246	
CEACAM1 (carcinoembryonic antigen-related cell adhesion molecule 1 (biliary glycoprotein))	X16354	Hs.50964	NM_001712	19q13.2	X16354 /FEATURE= /DEFINITION=HSTM1CEA Human mRNA for transmembrane carcinoembryonic antigen BGPa (formerly TM1-CEA)	988_at
GCLC (glutamate-cysteine ligase, catalytic subunit)	M90656	Hs.151393	NM_001498	6p12	Cluster Incl. M90656:Human gamma-glutamylcysteine synthetase (GCS) mRNA, complete cds /cds=(92,2005) /gb=M90656 /gi=183038 /ug=Hs.151393 /len=2615	31850_at
IGHM (immunoglobulin heavy constant mu)	X67301	Hs.302063		14q32.33	Cluster Incl. :H.sapiens mRNA for IgM heavy chain constant region (Ab63) /cds=(0,1361) /gb=X67301 /gi=38407 /ug=Hs.179543 /len=1453	41164_at
HIRA (HIR (histone cell cycle regulation defective, S. cerevisiae) homolog A)	X89887	Hs.172350	NM_003325	22q11.21	Cluster Incl. X89887:Homo sapiens mRNA for WD repeat protein (HIRA) /cds=(220,3273) /gb=X89887 /gi=3928218	32706_at

					/ug=Hs.172350 /len=4018	
EPOR (erythropoietin receptor)	M60459	Hs.89548	NM_000121	19p13.3-p13.2	M60459 /FEATURE= /DEFINITION=HUMERYTH Human erythropoietin receptor mRNA, complete cds	1087_at
LILRB2 (leukocyte immunoglobulin-like receptor, subfamily B (with TM and ITIM domains), member	AF004231	Hs.22405	NM_005874	19q13.4	Cluster Incl. AF004231:Homo sapiens monocyte/macrophage Ig-related receptor MIR-10 (MIR cl-10) mRNA, complete cds /cds=(208,2001) /gb=AF004231 /gi=2343110 /ug=Hs.22405 /len=2863	39221_at
KLF5 (Kruppel-like factor 5 (intestinal))	D14520	Hs.84728	NM_001730	13q21.2-13q22.2	Cluster Incl. D14520:Human mRNA for GC-Box binding protein BTEB2, complete cds /cds=(558,1217) /gb=D14520 /gi=303596 /ug=Hs.84728 /len=1301	37926_at
MTMR1 (myotubularin related protein 1)	AJ224979	Hs.23200		Xq28	Cluster Incl. AJ224979:Homo sapiens mRNA for MTMR1 protein /cds=(0,1990) /gb=AJ224979 /gi=4128155 /ug=Hs.23200	34654_at

					/len=2582	
KEL (Kell blood group)	M64934	Hs.157	NM_000420	7q33	Cluster Incl. M64934:Human kell blood group protein mRNA /cds=(123,2321) /gb=M64934 /gi=413776 /ug=Hs.157 /len=2458	38197_at
MAL (mal, T-cell differentiation protein)	X76220	Hs.80395	NM_002371	2cen-q13	Cluster Incl. X76220:H.sapiens MAL gene exon 1 (and joined CDS) /cds=(59,520) /gb=X76220 /gi=433225 /ug=Hs.80395 /len=1056	38051_at
CDC2 (cell division cycle 2, G1 to S and G2 to M)	D88357	Hs.184572	NM_001786	10q21.1	Cluster Incl. D88357:Homo sapiens mRNA for CDC2 delta T, complete cds /cds=(27,749) /gb=D88357 /gi=3126638 /ug=Hs.184572 /len=780	33324_s_at
IGLL3 (immunoglobulin lambda-like polypeptide 3)	AI932613	Hs.296552		22q11.23	Cluster Incl. AI932613:wo05c02.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-2454434 /clone_end=3 /gb=AI932613 /gi=5671350 /ug=Hs.62036 /len=570	41827_f_at

IL7R (interleukin 7 receptor)	M29696	Hs.237868	NM_002185	5p13	M29696 /FEATURE= /DEFINITION=HUMIL7AA Human interleukin-7 receptor (IL-7) mRNA, complete cds	1370_at
EGFL5 (EGF-like-domain, multiple 5)	AB011542	Hs.5599		9q32-q33.3	Cluster Incl. AB011542:Homo sapiens mRNA for MEGF9, partial cds /cds=(0,1129) /gb=AB011542 /gi=3449309 /ug=Hs.5599 /len=5507	36488_at
GW112(differentially expressed in hematopoietic lineages)	AF097021	Hs.273321	NM_006418	13	Cluster Incl. AF097021:Homo sapiens GW112 protein (GW112) mRNA, complete cds /cds=(508,1071) /gb=AF097021 /gi=3860076 /ug=Hs.100347 /len=2830	38615_at
NME2 (*non-metastatic cells 2, protein (NM23B) expressed in*)	X58965	Hs.275163	NM_002512	17q21.3	X58965 /FEATURE= /DEFINITION=HSNM23H2G H.sapiens RNA for nm23-H2 gene	1980_s_at
SET (SET translocation (myeloid leukemia-associated))	M93651	Hs.145279	NM_003011	9q34	Cluster Incl. M93651:Human set gene, complete cds /cds=(3,836) /gb=M93651 /gi=338038 /ug=Hs.145279 /len=2562	40189_at

					/gi=338038 /ug=Hs.145279 /len=2562	
KDELR1 (KDEL (Lys-Asp-Glu-Leu) endoplasmic reticulum protein retention receptor 1)	X55885	Hs.78040	NM_006801	19q13.3	Cluster Incl. X55885:Human mRNA for a presumptive KDEL receptor /cds=(146,784) /gb=X55885 /gi=34030 /ug=Hs.78040 /len=1086	37386_i_at
NME2 (non-metastatic cells 2, protein (NM23B) expressed in)	X58965	Hs.275163	NM_002512	17q21.3	Cluster Incl. X58965:H.sapiens RNA for nm23-H2 gene /cds=(72,530) /gb=X58965 /gi=35069 /ug=Hs.227823 /len=670	33415_at
KARS (lysyl-tRNA synthetase)	D32053	Hs.3100	NM_005548	16q23-q24	Cluster Incl. D32053:Homo sapiens mRNA for Lysyl tRNA Synthetase, complete cds /cds=(40,1833) /gb=D32053 /gi=2366751 /ug=Hs.3100 /len=1997	34336_at
(NOT approved by the HUGO/GDB nomenclature commitee)	U41635	Hs.76228	NM_006812	12	Cluster Incl. U41635:Human OS-9 precurosor mRNA, complete cds /cds=(85,2088) /gb=U41635 /gi=1322233 /ug=Hs.76228 /len=2738	36996_at

PLOD3 (procollagen-lysine, 2-oxoglutarate 5-dioxygenase 3)	AF046889	Hs.153357	NM_001084	7q22	Cluster Incl. AF046889:Homo sapiens lysyl hydroxylase isoform 3 (PLOD3) mRNA, complete cds /cds=(216,2432) /gb=AF046889 /gi=3153234 /ug=Hs.153357 /len=2735	39801_at
(NOT approved by the HUGO/GDB nomenclature commitee)	AC004410	Hs.284161		19	Cluster Incl. AC004410:Homo sapiens chromosome 19, fosmid 39554 /cds=(0,1196) /gb=AC004410 /gi=2959558 /ug=Hs.167352 /len=1197	35426_at
P2RX4 (purinergic receptor P2X, ligand-gated ion channel, 4)	U83993	Hs.321709	NM_002560	12q24.32	Cluster Incl. U83993:Human P2X4 purinoreceptor mRNA, complete cds /cds=(309,1475) /gb=U83993 /gi=4099120 /ug=Hs.9610 /len=2031	38332_at
COMT (catechol-O-methyltransferase)	M58525	Hs.240013	NM_000754	22q11.21	Cluster Incl. M58525:Homo sapiens catechol-O-methyltransferase (COMT) mRNA, complete cds /cds=(204,1019) /gb=M58525 /gi=179954 /ug=Hs.78534 /len=1206	34651_at

UQCRC2 (ubiquinol-cytochrome c reductase core protein II)	J04973	Hs.173554	NM_003366	16p12	Cluster Incl. J04973:Human cytochrome bc-1 complex core protein II mRNA, complete cds /cds=(53,1414) /gb=J04973 /gi=180927 /ug=Hs.173554 /len=1588	40854_at
EIF4A1 (eukaryotic translation initiation factor 4A, isoform 1)	D13748	Hs.129673	NM_001416	17p13	D13748 /FEATURE= /DEFINITION=HUM4AI Human mRNA for eukaryotic initiation factor 4AI	1199_at
(NOT approved by the HUGO/GDB nomenclature commitee)	AI582831	Hs.102419	NM_015871	1	Cluster Incl. AI582831:tn36c06.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-2169706 /clone_end=3 /gb=AI582831 /gi=4568728 /ug=Hs.102419 /len=555	38640_at
NFIL3 (nuclear factor, interleukin 3 regulated)	X64318	Hs.79334	NM_005384	9q22	Cluster Incl. X64318:H.sapiens E4BP4 gene /cds=(213,1601) /gb=X64318 /gi=30955 /ug=Hs.79334 /len=1904	37544_at

(NOT approved by the HUGO/GDB nomenclature commitee)	AL022328	Hs.33026	NM_025204	22	Cluster Incl. AL022328:Human DNA sequence from clone 402G11 on chromosome 22q13.31-13.33 Contains genes for SAPK3 (stress-activated protein kinase 3), PRKM11 (protein kinase mitogen-activated 11), KIAA0315, ESTs, GSSs and CpG islands /cds=(11,1105) /gb=AL	40033_at
GSTP1 (glutathione S-transferase pi)	U12472	Hs.226795	NM_000852	11q13	Cluster Incl. U12472:Human glutathione S-transferase (GST phi) gene, complete cds /cds=(0,632) /gb=U12472 /gi=763404 /ug=Hs.226795 /len=757	33396_at
GRHPR (glyoxylate reductase/hydroxypyruvate reductase)	W28944	Hs.155742	NM_012203	9q12	Cluster Incl. W28944:54h12 Homo sapiens cDNA /gb=W28944 /gi=1308955 /ug=Hs.155742 /len=748	40133_s_at
HSPCB (heat shock 90kD protein 1, beta)	M16660	Hs.74335	NM_007355	6p12	Cluster Incl. M16660:Human 90-kDa heat-shock protein gene, cDNA, complete cds /cds=UNKNOWN /gb=M16660 /gi=184420	33984_at

					/ug=Hs.74335 /len=2543	
(NOT approved by the HUGO/GDB nomenclature commitee)	U70671	Hs.43509	NM_007245	7	Cluster Incl. U70671:Human ataxin-2 related protein mRNA, partial cds /cds=(0,1044) /gb=U70671 /gi=1679685 /ug=Hs.43509 /len=1206	34817_s_at
ARHGDIA (Rho GDP dissociation inhibitor (GDI) alpha)	X69550	Hs.159161	NM_004309	17q25.3	Cluster Incl. X69550:H.sapiens mRNA for rho GDP-dissociation Inhibitor 1 /cds=(53,667) /gb=X69550 /gi=456190 /ug=Hs.159161 /len=1819	40164_at
(NOT approved by the HUGO/GDB nomenclature commitee)	S82470	Hs.78768	NM_024298	19	S82470 /FEATURE= /DEFINITION=S82470 BB1=malignant cell expression-enhanced gene/tumor progression-enhanced gene [human, UM-UC-9 bladder carcinoma cell line, mRNA, 1897 nt]	181_g_at
DF (D component of complement (adipsin))	M84526	Hs.155597	NM_001928	19	Cluster Incl. M84526:Human adipsin/complement factor D mRNA, complete cds /cds=(54,740) /gb=M84526	40282_s_at

					/gi=178625 /ug=Hs.155597 /len=1071	
CLN3 (ceroid-lipofuscinosis, neuronal 3, juvenile (Batten, Spielmeyer-Vogt disease))	AC002544	Hs.194660	NM_000086	16p12.1	Cluster Incl. AC002544:Homo sapiens Chromosome 16 BAC clone CIT987SK-A-761H5 /cds=(85,2826) /gb=AC002544 /gi=3337382 /ug=Hs.4835 /len=3027	34841_at
DAP (death-associated protein)	X76105	Hs.75189	NM_004394	5p15.2	Cluster Incl. X76105:H.sapiens DAP-1 mRNA /cds=(159,467) /gb=X76105 /gi=434844 /ug=Hs.75189 /len=2232	36199_at
NACA (nascent-polypeptide-associated complex alpha polypeptide)	AF054187	Hs.32916	NM_005594	12q23-q24.1	Cluster Incl. AF054187:Homo sapiens alpha NAC mRNA, complete cds /cds=(309,956) /gb=AF054187 /gi=4092059 /ug=Hs.146763 /len=1059	39740_g_at

Table 4:

UCL/HGNC/HUGO Human Gene Nomenclature Database Symbol	GenBank Accession No.	UniGene Cluster	RefSeq	Chromosomal Location	Description Unigene Build #95	Gene Name
KIAA0842(KIAA0842 protein)	AB020649	Hs.74569		1	Cluster Incl. AB020649:Homo sapiens mRNA for KIAA0842 protein, partial cds /cds=(0,3062) /gb=AB020649 /gi=4240172 /ug=Hs.74569 /len=3896	36150_at
H1FX (H1 histone family, member X)	D64142	Hs.109804	NM_006026		D64142 /FEATURE=mRNA /DEFINITION=D64142 Human mRNA for histone H1x, complete cds	319_g_at
NUFIP1 (nuclear fragile X mental retardation protein interacting protein 1)	AL049842	Hs.120247	NM_012345	13q14	Cluster Incl. AL049842:Human DNA sequence from clone 129L7 on chromosome 6q12-13. Contains the gene for a PUTATIVE novel protein, ESTs, an STS, GSSs and a taga repeat polymorphism /cds=(9,749) /gb=AL049842	37518_at

					/gi=5419768 /ug=Hs.120247 /len=1679	
PPP6C (protein phosphatase 6, catalytic subunit)	X92972	Hs.80324	NM_002721	xq22.3	Cluster Incl. X92972:H.sapiens mRNA for protein phosphatase 6 /cds=(21,938) /gb=X92972 /gi=5701862 /ug=Hs.80324 /len=1292	37581_at
SUI1(putative translation initiation factor)	AJ012375	Hs.150580	NM_005801	19	Cluster Incl. AJ012375:Homo sapiens mRNA for SUI1 protein translation initiation factor /cds=UNKNOWN /gb=AJ012375 /gi=4468342 /ug=Hs.150580 /len=1350	40203_at
UNRIP(unr-interacting protein)	AB024327	Hs.3727	NM_007178	12	Cluster Incl. AB024327:Homo sapiens pt-wd mRNA for WD-40 repeat protein, complete cds /cds=(300,1352) /gb=AB024327 /gi=4519416 /ug=Hs.3727 /len=1850	34402_at

ATF4 (activating transcription factor 4 (tax-responsive enhancer element B67))	AL022312	Hs.181243	NM_001675	22q13.1	Cluster Incl. AL022312:dJ1104E15.2 (activating transcription factor 4 (tax-responsive enhancer element B67)) /cds=(882,1937) /gb=AL022312 /gi=4914501 /ug=Hs.181243 /len=2016	41235_at
WBSCR1 (Williams-Beuren syndrome chromosome region 1)	D26068	Hs.180900	NM_022170	7q11.23	Cluster Incl. D26068:Human mRNA for KIAA0038 gene, partial cds /cds=(0,694) /gb=D26068 /gi=436225 /ug=Hs.180900 /len=2477	41212_r_at
RNPS1 (RNA-binding protein S1, serine-rich domain)	L37368	Hs.75104	NM_006711	16p13.3	Cluster Incl. L37368:Human (clone E5.1) RNA-binding protein mRNA, complete cds /cds=(549,1466) /gb=L37368 /gi=1236282 /ug=Hs.75104 /len=2438	36186_at
C6orf5 (chromosome 6 open reading frame 5)	AL050289	Hs.7446	NM_015524	6q21	Cluster Incl. AL050289:Homo sapiens mRNA; cDNA DKFZp586G0522 (from clone DKFZp586G0522) /cds=(179,1876) /gb=AL050289 /gi=4886510 /ug=Hs.7446 /len=2364	36139_at

	W28807				Cluster Incl. W28807:52a3 Homo sapiens cDNA /gb=W28807 /gi=1308755 /ug=Hs.121849 /len=819	39370_at
POLB (polymerase (DNA directed), beta)	D29013	Hs.180107	NM_002690	8p11.2	D29013 /FEATURE= /DEFINITION=HUMLNCAP Human mRNA for DNA polymerase beta, complete cds	1696_at
DKFZP434D1335(DKFZP434D1335 protein)	AI920820	Hs.8258		19	Cluster Incl. AI920820:wn82e10.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-2452362 /clone_end=3 /gb=AI920820 /gi=5656784 /ug=Hs.8258 /len=519	38400_at
PRCC (papillary renal cell carcinoma (translocation-associated))	X99720	Hs.9629	NM_005973	1q21.1	Cluster Incl. X99720:H.sapiens TPRC gene /cds=(212,1687) /gb=X99720 /gi=1869817 /ug=Hs.9629 /len=2053	39149_at
GNA15 (guanine nucleotide binding protein (G protein), alpha 15 (Gq class))	M63904	Hs.73797	NM_002068	19p13.3	Cluster Incl. M63904:Human G-alpha 16 protein mRNA, complete cds /cds=(219,1343) /gb=M63904 /gi=182891 /ug=Hs.73797 /len=2060	40365_at

RBM4 (RNA binding motif protein 4)	U89505	Hs.6106	NM_002896	11q13	Cluster Incl. U89505:Human Hlark mRNA, complete cds /cds=(55,1155) /gb=U89505 /gi=2078528 /ug=Hs.6106 /len=1598	35351_at
TCEB3 (transcription elongation factor B (SIII), polypeptide 3 (110kD, elongin A))	L47345	Hs.155202	NM_003198	1p36.1	L47345 /FEATURE= /DEFINITION=HUMELONA Homo sapiens elongin A mRNA, complete cds	639_s_at
BLCAP (bladder cancer associated protein)	AL049288	Hs.5300	NM_006698	20q11.2-q12	Cluster Incl. AL049288:Homo sapiens mRNA; cDNA DKFZp564M053 (from clone DKFZp564M053) /cds=UNKNOWN /gb=AL049288 /gi=4500049 /ug=Hs.5300 /len=2018	35267_g_at
SNRPA1 (small nuclear ribonucleoprotein polypeptide A')	X13482	Hs.80506	NM_003090	22q	Cluster Incl. X13482:Human mRNA for U2 snRNP-specific A protein /cds=(56,823) /gb=X13482 /gi=37546 /ug=Hs.80506 /len=1033	37585_at
ZNF207 (zinc finger protein 207)	AF046001	Hs.62112	NM_003457	17p13.2	Cluster Incl. AF046001:Homo sapiens zinc finger transcription factor (ZNF207) mRNA, complete cds /cds=(202,1638)	35368_at

					/gb=AF046001 /gi=2895869 /ug=Hs.62112 /len=2331	
KIAA0174(KIAA0174 gene product)	D79996	Hs.75824	NM_014761	19	Cluster Incl. D79996:Human mRNA for KIAA0174 gene, complete cds /cds=(63,1157) /gb=D79996 /gi=1136407 /ug=Hs.75824 /len=2348	36942_at
UBE2D3 (ubiquitin-conjugating enzyme E2D 3 (homologous to yeast UBC4/5)	U39318	Hs.118797	NM_003340	4q24-q26	Cluster Incl. U39318:Human E2 ubiquitin conjugating enzyme UbcH5C (UBCH5C) mRNA, complete cds /cds=(45,488) /gb=U39318 /gi=1145690 /ug=Hs.118797 /len=724	39083_at
H1FX (H1 histone family, member X)	D64142	Hs.109804	NM_006026		D64142 /FEATURE=mRNA /DEFINITION=D64142 Human mRNA for histone H1x, complete cds	318_at
KIF5A (kinesin family member 5A)	U06698	Hs.192760	NM_004984	12q13	Cluster Incl. U06698:Human neuronal kinesin heavy chain mRNA, complete cds /cds=(148,3246) /gb=U06698 /gi=497123	35880_at

					/ug=Hs.192760 /len=3840	
ELF1 (E74-like factor 1 (ets domain transcription factor)	M82882	Hs.154365		13q13	Cluster Incl. M82882:Human cis-acting sequence /cds=UNKNOWN /gb=M82882 /gi=180551 /ug=Hs.154365 /len=3564	40067_at
ZNF278 (zinc finger protein 278)	AI352450	Hs.27801	NM_014323	22q12.2	Cluster Incl. AI352450:qt16g11.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-1947812 /clone_end=3 /gb=AI352450 /gi=4089656 /ug=Hs.27801 /len=508	34300_at
TCF7L2 (transcription factor 7-like 2 (T-cell specific, HMG-box))	Y11306	Hs.173638	NM_030756	10q25.3	Cluster Incl. Y11306:Homo sapiens mRNA for hTCF-4 /cds=(307,2097) /gb=Y11306 /gi=4469251 /ug=Hs.154485 /len=2444	32025_at
RNF10 (ring finger protein 10)	D87451	Hs.5094	NM_014868	12	Cluster Incl. D87451:Human mRNA for KIAA0262 gene, complete cds /cds=(698,2983) /gb=D87451 /gi=1665790 /ug=Hs.5094 /len=3205	34883_at

TACC1 (transforming, acidic coiled-coil containing protein 1)	AF049910	Hs.173159	NM_006283	8p11	Cluster Incl. AF049910:Homo sapiens TACC1 (TACC1) mRNA, complete cds /cds=(320,2737) /gb=AF049910 /gi=3435156 /ug=Hs.173159 /len=7735	40841_at
PTS (6-pyruvoyltetrahydropterin synthase)	L76259	Hs.366	NM_000317	11q22.3-q23.3	Cluster Incl. L76259:Homo sapiens PTS gene, complete cds /cds=(68,505) /gb=L76259 /gi=2276403 /ug=Hs.366 /len=921	35697_at
MT1A (metallothionein 1A (functional))	K01383	Hs.173451		16q13	Cluster Incl. K01383:Human metallothionein-I-A gene, complete coding sequence /cds=(0,185) /gb=K01383 /gi=187536 /ug=Hs.203967 /len=186	31623_f_at
CAMK2G (calcium/calmodulin-dependent protein kinase (CaM kinase) II gamma)	U81554	Hs.250857		10q22	Cluster Incl. U81554:Homo sapiens CaM kinase II isoform mRNA, complete cds /cds=(212,931) /gb=U81554 /gi=2275253 /ug=Hs.231812 /len=972	31670_s_at
SFPQ (splicing factor proline/glutamine rich (polypyrimidine tract-binding protein-	W27050	Hs.180610	NM_005066	1pter-p32.3	Cluster Incl. W27050:19f7 Homo sapiens cDNA /gb=W27050 /gi=1306422	41199_s_at

associated))					/ug=Hs.180610 /len=699	
PPP2CA (protein phosphatase 2 (formerly 2A), catalytic subunit, alpha isoform)	M60483	Hs.91773	NM_002715	5q23-q31	M60483 /FEATURE=mRNA /DEFINITION=HUMPP2AA Human protein phosphatase 2A catalytic subunit-alpha gene, complete cds	237_s_at
KIAA0105(Wilms' tumour 1-associating protein	D14661	Hs.119	NM_004906	6	Cluster Incl. D14661:Human mRNA for KIAA0105 gene, complete cds /cds=(124,579) /gb=D14661 /gi=285946 /ug=Hs.119 /len=1622	41635_at
GNA13 (guanine nucleotide binding protein (G protein), alpha 13)	L22075	Hs.1666	NM_006572	17q22-q24	Cluster Incl. L22075:Human guanine nucleotide regulatory protein (G13) mRNA, complete cds /cds=(41,1174) /gb=L22075 /gi=404721 /ug=Hs.1666 /len=1402	33635_at
UGTREL7(UDP-glucuronic acid/UDP-N-acetylgalactosamine dual transporter)	D87449	Hs.82635	NM_015139	1	Cluster Incl. D87449:Human mRNA for KIAA0260 gene, partial cds /cds=(0,1153) /gb=D87449 /gi=1665786 /ug=Hs.82635 /len=5918	37888_at

BRAP (BRCA1 associated protein)	AL042733	Hs.122764	NM_006768	12q24	Cluster Incl. AL042733:DKFZp434B2222_s1 Homo sapiens cDNA, 3 end /clone=DKFZp434B2222 /clone_end=3 /gb=AL042733 /gi=5422182 /ug=Hs.30982 /len=782	41512_at
PMAIP1 (phorbol-12-myristate-13-acetate-induced protein 1)	D90070	Hs.96	NM_021127	18q22	Cluster Incl. D90070:Human ATL-derived PMA-responsive (APR) peptide mRNA /cds=(173,337) /gb=D90070 /gi=219475 /ug=Hs.96 /len=1885	41048_at
KIAA1116(KIAA1116 protein)	AB029039	Hs.227602	NM_014892	6	Cluster Incl. AB029039:Homo sapiens mRNA for KIAA1116 protein, complete cds /cds=(185,4000) /gb=AB029039 /gi=5689568 /ug=Hs.227602 /len=4664	34274_at
ERF (Ets2 repressor factor)	U15655	Hs.333069	NM_006494	19q13	U15655 /FEATURE= /DEFINITION=HSU15655 Human ets domain protein ERF mRNA, complete cds	1242_at

ZNF161 (zinc finger protein 161)	D28118	Hs.6557	NM_007146	3q26.2	D28118 /FEATURE= /DEFINITION=HUMDB1 Human mRNA for DB1, complete cds	350_at
MAPK3 (mitogen-activated protein kinase 3)	X60188	Hs.861		16p12-p11.2	X60188 /FEATURE=mRNA /DEFINITION=HSERK1 Human ERK1 mRNA for protein serine/threonine kinase	1000_at
	AI659108				Cluster Incl. AI659108:tu08c09.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-2250448 /clone_end=3 /gb=AI659108 /gi=4762678 /ug=Hs.99093 /len=492	41084_at
SYPL (synaptophysin-like protein)	X68194	Hs.80919	NM_006754	7q22.1-7q31.33	Cluster Incl. X68194:H.sapiens h-Sp1 mRNA /cds=(33,812) /gb=X68194 /gi=32473 /ug=Hs.80919 /len=2108	38075_at
SUPT5H (suppressor of Ty (S.cerevisiae) 5 homolog)	AF040253	Hs.70186	NM_003169	19q13	Cluster Incl. AF040253:Homo sapiens transcription factor Tat-CT1 mRNA, complete cds /cds=(207,3470) /gb=AF040253 /gi=4104823 /ug=Hs.70186	35826_at

					/len=3738	
D13S106E(highly charged protein)	X59131	Hs.151236	NM_005800	13	Cluster Incl. X59131:Homo sapiens D13S106 mRNA for a highly charged amino acid sequence /cds=(177,3455) /gb=X59131 /gi=3776087 /ug=Hs.151236 /len=3650	31847_at
WHN (winged-helix nude)	Y11739	Hs.198313	NM_003593	7q11-q12	Cluster Incl. Y11739:H.sapiens mRNA for whn transcription factor /cds=(29,1975) /gb=Y11739 /gi=2315191 /ug=Hs.198313 /len=2697	31980_at
ARHA (ras homolog gene family, member A)	L25080	Hs.77273	NM_001664	3p21.3	L25080 /FEATURE= /DEFINITION=HUMRHOAA Homo sapiens GTP-binding protein (rhoA) mRNA, complete cds	1394_at
ERF (Ets2 repressor factor)	U15655	Hs.333069	NM_006494	19q13	Cluster Incl. U15655:Human ets domain protein ERF mRNA, complete cds /cds=(122,1768) /gb=U15655 /gi=1015336	38996_at

					/ug=Hs.110906 /len=2667	
MAX (MAX protein)	X60287	Hs.42712	NM_002382	14q23	X60287 /FEATURE=cds /DEFINITION=HSMAXM H.sapiens max mRNA	1981_s_at
RAB14(GTPase Rab14)	AF052113	Hs.5807	NM_016322	9	Cluster Incl. AF052113:Homo sapiens clone 23675 mRNA sequence /cds=UNKNOWN /gb=AF052113 /gi=3360420 /ug=Hs.5807 /len=1652	35325_at
EIF3S5 (eukaryotic translation initiation factor 3, subunit 5 (epsilon, 47kD))	U94855	Hs.7811	NM_003754	2p16.1	Cluster Incl. U94855:Homo sapiens translation initiation factor 3 47 kDa subunit mRNA, compl U94855 cds /cds=(6,1079) /gb=U94855 /gi=2055430 /ug=Hs.7811 /len=1231	32576_at
DNAJA1 (DnaJ (Hsp40) homolog, subfamily A, member 1)	L08069	Hs.94	NM_001539	9p13-p12	Cluster Incl. L08069:Human heat shock protein, E. coli DnaJ homologue mRNA, complete cds /cds=(82,1275) /gb=L08069 /gi=306713 /ug=Hs.94 /len=1438	39118_at

SLBP (stem-loop (histone) binding protein)	U75679	Hs.75257	NM_006527	4p16.3	Cluster Incl. U75679:Human histone stem-loop binding protein (SLBP) mRNA, complete cds /cds=(115,927) /gb=U75679 /gi=1732076 /ug=Hs.75257 /len=1725	36913_at
CCNI (cyclin I)	D50310	Hs.79933	NM_006835	4	D50310 /FEATURE= /DEFINITION=HUMCYI Human mRNA for cyclin I, complete cds	1836_at
	L10910				Cluster Incl. L10910:Homo sapiens splicing factor (CC1.3) mRNA, complete cds /cds=(149,1723) /gb=L10910 /gi=405191 /ug=Hs.145696 /len=2595	39725_at
SSH3BP1 (spectrin SH3 domain binding protein 1)	AF001628	Hs.24752	NM_005470	10p11.2	Cluster Incl. AF001628:Homo sapiens interactor protein AbIBP4 (AbIBP4) mRNA, complete cds /cds=(48,1403) /gb=AF001628 /gi=4100618 /ug=Hs.204036 /len=2175	38924_s_at
DYRK1A (dual-specificity tyrosine-(Y)-phosphorylation regulated kinase 1A	D86550	Hs.75842	NM_001396	21q22.13	D86550 /FEATURE= /DEFINITION=D86550 Human mRNA for	1512_at

phosphorylation regulated kinase 1A					serine/threonine protein kinase, complete cds	
ZFR(zinc finger RNA binding protein)	AI459274	Hs.173518	NM_016107	5	Cluster Incl. AI459274:tk11f11.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-2150733 /clone_end=3 /gb=AI459274 /gi=4311853 /ug=Hs.87150 /len=687	33148_at
BAG5 (BCL2-associated athanogene 5)	AB020680	Hs.5443	NM_004873	14	Cluster Incl. AB020680:Homo sapiens mRNA for KIAA0873 protein, partial cds /cds=(0,1400) /gb=AB020680 /gi=4240234 /ug=Hs.5443 /len=4119	36463_at
GABPB1 (*GA-binding protein transcription* factor, beta subunit 1 (53kD))	D13317	Hs.78915	NM_005254	7q11.2	Cluster Incl. D13317:Human mRNA for transcription factor, E4TF1-53, complete cds /cds=(205,1356) /gb=D13317 /gi=286024 /ug=Hs.211616 /len=1552	35943_s_at
PTP4A2 (protein tyrosine phosphatase type IVA, member 2)	U14603	Hs.82911	NM_003479	1p35	Cluster Incl. U14603:Human protein-tyrosine phosphatase (HU-PP-1) mRNA, partial sequence /cds=(423,926) /gb=U14603 /gi=894158 /ug=Hs.82911	38415_at

					/len=1526	
KIAA0414(KIAA0414 protein)	AB007874			9	Cluster Incl. AB007874:Homo sapiens KIAA0414 mRNA, partial cds /cds=(1132,2535) /gb=AB007874 /gi=2887448 /ug=Hs.127649 /len=5725	41695_at
DLST (dihydrolipoamide S-succinyltransferase (E2 component of 2-oxo-glutarate complex))	D26535	Hs.250801	NM_001933	14q24.3	Cluster Incl. D26535:Human gene for dihydrolipoamide succinyltransferase, complete cds (exon 1-15) /cds=(43,1404) /gb=D26535 /gi=537349 /ug=Hs.179989 /len=2822	33258_g_at
KIAA0016(translocase of outer mitochondrial membrane 20 (yeast) homolog)	D13641	Hs.75187	Hs.75187	1	Cluster Incl. D13641:Human mRNA for KIAA0016 gene, complete cds /cds=(101,538) /gb=D13641 /gi=285986 /ug=Hs.75187 /len=3259	36198_at
HT010(uncharacterized hypothalamus protein HT010	AL049948	Hs.6375	NM_018471	2	Cluster Incl. AL049948:Homo sapiens mRNA; cDNA DKFZp564K0222 (from clone DKFZp564K0222) /cds=UNKNOWN	35750_at

					/gb=AL049948 /gi=4884195 /ug=Hs.6375 /len=1027	
CACNA1E (calcium channel, voltage-dependent, alpha 1E subunit)	L29385	Hs.166110	NM_000721	1q25-q31	Cluster Incl. L29385:Homo sapiens (clone pcDNA-alpha1E-3) voltage-dependent calcium channel alpha-1E-3 subunit mRNA, complete cds /cds=(165,6977) /gb=L29385 /gi=495869 /ug=Hs.166110 /len=7089	33624_at
ZFX (zinc finger protein, X-linked)	X59739	Hs.2074	NM_003410	xp21.3	Cluster Incl. X59739:Human ZFX mRNA for put. transcription activator, isoform 2 /cds=(78,2492) /gb=X59739 /gi=38021 /ug=Hs.2074 /len=5527	38931_at
TIAL1 (TIA1 cytotoxic granule-associated RNA-binding protein-like 1)	M96954	Hs.182741	NM_022333	10q	Cluster Incl. M96954:Homo sapiens nucleolysin TIAR mRNA, complete cds /cds=(45,1172) /gb=M96954 /gi=307313 /ug=Hs.182741 /len=1401	41761_at
	L03426				Cluster Incl. L03426:Human XE7 mRNA, complete alternate coding regions	34215_at

					/cds=(166,1323) /gb=L03426 /gi=340386 /ug=Hs.21595 /len=3233	
EP300 (E1A binding protein p300)	U01877	Hs.25272	NM_001429	22q13.2	Cluster Incl. U01877:Human p300 protein mRNA, complete cds /cds=(1199,8443) /gb=U01877 /gi=495300 /ug=Hs.25272 /len=9046	33896_at
RLBP1 (retinaldehyde-binding protein 1)	L34219	Hs.1933	NM_000326	15q26	Cluster Incl. L34219:Homo sapiens retinaldehyde-binding protein (CRALBP) gene, complete cds /cds=(100,1053) /gb=L34219 /gi=598228 /ug=Hs.1933 /len=1450	35887_at
ERAL1 (Era (E. coli G-protein homolog)-like 1)	AF082657	Hs.3426		17q11.2	Cluster Incl. AF082657:Homo sapiens Era GTPase A protein (HERA-A) mRNA, partial cds /cds=(0,1332) /gb=AF082657 /gi=3415108 /ug=Hs.3426 /len=1839	34379_at
RAC1 (ras-related C3 botulinum toxin substrate 1 (rho family, small GTP binding protein Rac1))	D25274	Hs.173737	NM_018890	7p22	Cluster Incl. D25274:Homo sapiens mRNA, clone-PO2ST9 /cds=UNKNOWN /gb=D25274 /gi=464185 /ug=Hs.173737	40864_at

					/len=1232	
HPRP4P(PRP4/STK/WD splicing factor)	AI184802	Hs.8551	NM_004697	9	Cluster Incl. AI184802:qd24g04.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-1724694 /clone_end=3 /gb=AI184802 /gi=3735440 /ug=Hs.8551 /len=625	37938_at
KIN (antigenic determinant of recA protein (mouse) homolog)	AJ005273	Hs.123647	NM_012311	10p15-p14	Cluster Incl. AJ005273:Homo sapiens mRNA for Kin17 protein /cds=(65,1246) /gb=AJ005273 /gi=3850703 /ug=Hs.123647 /len=1518	37778_at
P2RX4 (purinergic receptor P2X, ligand-gated ion channel, 4)	U83993	Hs.321709	NM_002560	12q24.32	Cluster Incl. U83993:Human P2X4 purinoreceptor mRNA, complete cds /cds=(309,1475) /gb=U83993 /gi=4099120 /ug=Hs.9610 /len=2031	38332_at
VDAC2 (voltage-dependent anion channel 2)	L08666	Hs.78902	NM_003375	10q22	Cluster Incl. L08666:Homo sapiens porin (por) mRNA, complete cds and truncated cds /cds=UNKNOWN /gb=L08666 /gi=190199 /ug=Hs.78902 /len=1464	37697_s_at

DSIPI (delta sleep inducing peptide, immunoreactor)	AI635895	Hs.75450	NM_004089	xp21.1-q25	Cluster Incl. AI635895:tz82a07.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-2295060 /clone_end=3 /gb=AI635895 /gi=4687225 /ug=Hs.75450 /len=1082	36629_at
ZNF134 (zinc finger protein 134 (clone pHZ-15))	U09412	Hs.357	NM_003435	19q13.4	Cluster Incl. U09412:Human zinc finger protein ZNF134 mRNA, complete cds /cds=(521,1567) /gb=U09412 /gi=488552 /ug=Hs.357 /len=2094	36295_at
SGT (small glutamine-rich tetratricopeptide repeat (TPR)-containing)	AL050156	Hs.203910	NM_003021	19p13	Cluster Incl. AL050156:Homo sapiens mRNA; cDNA DKFZp586N1020 (from clone DKFZp586N1020) /cds=(0,1050) /gb=AL050156 /gi=4884157 /ug=Hs.203910 /len=2277	32816_at
SRP19 (signal recognition particle 19kD)	X12791	Hs.2943	NM_003135	5q21-q22	Cluster Incl. X12791:Human mRNA for 19kD protein of signal recognition particle (SRP) /cds=(81,515) /gb=X12791 /gi=36112 /ug=Hs.2943 /len=894	35231_at

	W28360				Cluster Incl. W28360:46f9 Homo sapiens cDNA /gb=W28360 /gi=1308371 /ug=Hs.11498 /len=675	39040_at
MYCBP (c-myc binding protein)	D50692	Hs.78221	NM_012333	1p33-p32.2	D50692 /FEATURE= /DEFINITION=HUMAMY1 Homo sapiens mRNA for c-myc binding protein, complete cds	1904_at
KCNH2 (potassium voltage-gated channel, subfamily H (eag-related), member 2)	AF052728	Hs.188021	NM_000238	7q35-q36	Cluster Incl. AF052728:Homo sapiens HERG-USO (HERG) mRNA, alternatively spliced, partial cds /cds=(0,284) /gb=AF052728 /gi=3549258 /ug=Hs.165664 /len=767	38225_at
HMG4 (high-mobility group (nonhistone chromosomal) protein 4)	AL034450	Hs.19114	NM_005342	xq28	Cluster Incl. AL034450:Human DNA sequence from clone 115K14 on chromosome Xq22.3-23 Contains high mobility group protein 2a, ESTs, STS /cds=(0,605) /gb=AL034450 /gi=4210359 /ug=Hs.194749 /len=730	31588_at

TLN1 (talin 1)	AB028950	Hs.18420	NM_006289	9p13	Cluster Incl. AB028950:Homo sapiens mRNA for KIAA1027 protein, partial cds /cds=(0,5088) /gb=AB028950 /gi=5689390 /ug=Hs.18420 /len=5542	32166_at
PPP1R11 (protein phosphatase 1, regulatory (inhibitor) subunit 11)	U53588	Hs.82887	NM_021959	6p21.3	Cluster Incl. U53588:Homo sapiens MHC class 1 region /cds=(199,579) /gb=U53588 /gi=1685104 /ug=Hs.82887 /len=1607	38412_at
ature database symbol	S82297	Hs.75415	NM_004048	15q21-q22.2	S82297 /FEATURE= /DEFINITION=S82297 beta 2-microglobulin {11bp deleted between nucleotides 98-99} [human, colon cancer cell line HCT, mRNA Mutant, 416 nt]	201_s_at
ASH2L (ash2 (absent, small, or homeotic, Drosophila, homolog)-like)	AB022785	Hs.6856	NM_004674	8p11.2	Cluster Incl. AB022785:Homo sapiens ASH2L gene, complete cds, similar to Drosophila ash2 gene /cds=(12,1898) /gb=AB022785 /gi=4210446 /ug=Hs.6856 /len=2369	35804_at

RBMS1 (RNA binding motif, single stranded *interacting protein 1*)	Mssp-1	Hs.241567	NM_016839	2p14-q14.3	Single-Stranded Dna-Binding Protein Mssp-1	333_s_at
C3F(putative protein similar to nessy (Drosophila))	U72515	Hs.300423	NM_005768	12	Cluster Incl. U72515:Human C3f mRNA, complete cds /cds=(117,1262) /gb=U72515 /gi=1673519 /ug=Hs.189583 /len=1842	33710_at
TPM1 (tropomyosin 1 (alpha))	M19267	Hs.77899	NM_000366	15q22.1	Cluster Incl. M19267:Human tropomyosin mRNA, complete cds /cds=(286,1140) /gb=M19267 /gi=339943 /ug=Hs.77899 /len=1633	36791_g_at
LOC94392(hypothetical gene supported by AB007931; AF055010; AK001233; AK022322; AK022573; AK022924; AK023826; AK025149; AL049972; BC007962	AB007931			1	Cluster Incl. AB007931:Homo sapiens mRNA for KIAA0462 protein, partial cds /cds=(0,6831) /gb=AB007931 /gi=3413885 /ug=Hs.239686 /len=7150	33860_at
HSD17B4 (hydroxysteroid (17-beta) dehydrogenase 4)	X87176	Hs.75441	NM_000414	5q21	Cluster Incl. X87176:H.sapiens mRNA for 17-beta-hydroxysteroid dehydrogenase /cds=(48,2258) /gb=X87176 /gi=1050516	36626_at

					/ug=Hs.75441 /len=2593	
KIAA0728(KIAA0728 protein)	AB018271	Hs.198689		6	Cluster Incl. AB018271:Homo sapiens mRNA for KIAA0728 protein, partial cds /cds=(0,3197) /gb=AB018271 /gi=3882176 /ug=Hs.198689 /len=3964	32780_at
SR-BP1(sigma receptor (SR31747 binding protein 1))]	U79528	Hs.24447	NM_005866	9	Cluster Incl. U79528:Human SR31747 binding protein 1 mRNA, complete cds /cds=(74,745) /gb=U79528 /gi=1916799 /ug=Hs.24447 /len=1650	33879_at
CSPG2 (chondroitin sulfate proteoglycan 2 (versican))	X15998	Hs.81800	NM_004385	5q14.3	Cluster Incl. X15998:H.sapiens mRNA for the chondroitin sulphate proteoglycan versican, V1 splice-variant; precursor peptide /cds=(266,7495) /gb=X15998 /gi=37662 /ug=Hs.81800 /len=8224	38111_at
HCK (hemopoietic cell kinase)	M16591	Hs.89555	NM_002110	20q11-q12	Cluster Incl. M16591:Human hemopoietic cell protein-tyrosine kinase (HCK) gene, complete cds, clone lambda-a2/1a /cds=(168,1685) /gb=M16591 /gi=183911	40742_at

					/ug=Hs.89555 /len=2015	
	X92997				Cluster Incl. X92997:H.sapiens mRNA for IgG lambda light chain V-J-C region (clone Tgl4) /cds=(0,321) /gb=X92997 /gi=1070337 /ug=Hs.129722 /len=322	35530_f_at
TST (thiosulfate sulfurtransferase (rhodanese))	D87292	Hs.248267	NM_003312	22q13.1	Cluster Incl. D87292:Homo sapiens mRNA for rhodanese, complete cds /cds=(48,941) /gb=D87292 /gi=1877030 /ug=Hs.74097 /len=1137	38123_at
SSTR3 (somatostatin receptor 3)	M96738	Hs.225995	NM_001051	22q13.1	M96738 /FEATURE=cds /DEFINITION=HUMSSTR3X Human somatostatin receptor subtype 3 (SSTR3) gene, complete cds	557_s_at
KCNH2 (potassium voltage-gated channel, subfamily H (eag-related), member 2)	U04270	Hs.188021	NM_000238	7q35-q36	Cluster Incl. U04270:Human putative potassium channel subunit (h-erg) mRNA, complete cds /cds=(183,3662) /gb=U04270 /gi=487737 /ug=Hs.188021	38858_at

					/len=4070	
PTMS (parathymosin)	M24398	Hs.171814	NM_002824	17q12-q22	Cluster Incl. M24398:Human parathymosin mRNA, complete cds /cds=(300,608) /gb=M24398 /gi=339698 /ug=Hs.171814 /len=1109	40580_r_at
KIAA0859(KIAA0859 protein)]	AL049669	Hs.19469		1	Cluster Incl. AL049669:Human gene from PAC 612B18, chromosome 1 /cds=(272,1903) /gb=AL049669 /gi=4678746 /ug=Hs.19469 /len=2862	32262_at
SLC31A1 (solute carrier family 31 (copper transporters), member 1	U83460	Hs.73614	NM_001859	9q31-q32	Cluster Incl. U83460:Human high-affinity copper uptake protein (hCTR1) mRNA, complete cds /cds=(152,724) /gb=U83460 /gi=2315986 /ug=Hs.73614 /len=1804	40364_at
HSPD1 (heat shock 60kD protein 1 (chaperonin))	W28589	Hs.79037	NM_002156	12	Cluster Incl. W28589:48h12 Homo sapiens cDNA /gb=W28589 /gi=1308537 /ug=Hs.184567 /len=965	40913_at

PIK3CA (phosphoinositide-3-kinase, catalytic, alpha polypeptide)	P110	Hs.85701	NM_006218	3q26.3	Phosphatidylinositol 3-Kinase P110, Beta Isoform	1163_at
[KIAA1055(KIAA1055 protein)	AB028978	Hs.126084		15	Cluster Incl. AB028978:Homo sapiens mRNA for KIAA1055 protein, partial cds /cds=(0,2607) /gb=AB028978 /gi=5689446 /ug=Hs.126084 /len=5876	39400_at
WDR7 (WD repeat domain 7)	AB011113	Hs.10881		18q21.1-q22	Cluster Incl. AB011113:Homo sapiens mRNA for KIAA0541 protein, partial cds /cds=(0,3484) /gb=AB011113 /gi=3043605 /ug=Hs.10881 /len=6072	41430_at
PRKWNK1 (protein kinase, lysine deficient 1)	U00946	Hs.184592	NM_018979	12p13.3	Cluster Incl. U00946:Human clone A9A2BRB5 (CAC)n/(GTG)n repeat-containing mRNA /cds=UNKNOWN /gb=U00946 /gi=405048 /ug=Hs.184592 /len=1971	32185_at
TTC1 (tetratricopeptide repeat domain 1)	U46570	Hs.7733	NM_003314	5q32-q33.2	Cluster Incl. U46570:Human tetratricopeptide repeat protein (tpr1) mRNA, complete cds /cds=(50,928)	37321_at

					/gb=U46570 /gi=1688073 /ug=Hs.7733 /len=1407	
ST7 (suppression of tumorigenicity 7)	W02490	Hs.301974	NM_013437	8q22.2-q23.1	Cluster Incl. W02490:za48b02.r1 Homo sapiens cDNA, 5 end /clone=IMAGE-295755 /clone_end=5 /gb=W02490 /gi=1274488 /ug=Hs.5814 /len=623	40039_g_at
MINPP1 (multiple inositol polyphosphate histidine phosphatase, 1)	AL050356	Hs.95907	NM_004897	10q23	Cluster Incl. AL050356:Homo sapiens mRNA; cDNA DKFZp564L2016 (from clone DKFZp564L2016) /cds=UNKNOWN /gb=AL050356 /gi=4914568 /ug=Hs.95907 /len=2396	38325_at
NFYC (nuclear transcription factor Y, gamma)	Z74792	Hs.168157	NM_014223	1p32	Cluster Incl. Z74792:H.sapiens mRNA for CCAAT transcription binding factor subunit gamma /cds=(185,1192) /gb=Z74792 /gi=2564241 /ug=Hs.168157 /len=1965	40466_at
PK428(Ser-Thr protein kinase related to the myotonic dystrophy protein kinase)	U59305	Hs.44708	NM_003607	1	Cluster Incl. U59305:Human ser-thr protein kinase PK428 mRNA, complete cds /cds=(1288,2778) /gb=U59305	39962_at

					/gi=1695872 /ug=Hs.44708 /len=2785	
KIAA0674(KIAA0674 protein)]	AB014574	Hs.14799		9	Cluster Incl. AB014574:Homo sapiens mRNA for KIAA0674 protein, partial cds /cds=(0,3704) /gb=AB014574 /gi=3327161 /ug=Hs.14799 /len=4263	31826_at
IGKC (immunoglobulin kappa constant)	M63438	Hs.156110		2p12	Cluster Incl. M63438:Human Ig rearranged gamma chain mRNA, V-J-C region and complete cds /cds=(0,1049) /gb=M63438 /gi=184847 /ug=Hs.156110 /len=1244	38194_s_at
SEC22L1 (SEC22, vesicle trafficking protein (S. cerevisiae)-like 1	AF047442	Hs.50785	NM_004892	1q21.2-q21.3	Cluster Incl. AF047442:Homo sapiens vesicle trafficking protein sec22b mRNA, complete cds /cds=(64,711) /gb=AF047442 /gi=3335139 /ug=Hs.50785 /len=1433	41597_s_at
HCK (hemopoietic cell kinase)	M16592	Hs.89555	NM_002110	20q11-q12	M16592 /FEATURE=mRNA /DEFINITION=HUMHCKB Human hemopoietic cell protein-tyrosine kinase	2045_s_at

					(HCK) gene, complete cds, clone HK24	
TFDP1 (transcription factor Dp-1)	L23959	Hs.79353	NM_007111	13q34	Cluster Incl. L23959:Homo sapiens E2F-related transcription factor (DP-1) mRNA, complete cds /cds=(37,1269) /gb=L23959 /gi=414316 /ug=Hs.79353 /len=1440	37757_at
PCMT1 (protein-L-isoaspartate (D-aspartate) O-methyltransferase)	D13892	Hs.79137	NM_005389	6q24-q25	Cluster Incl. D13892:Human mRNA for carboxyl methyltransferase, complete cds /cds=(150,836) /gb=D13892 /gi=474983 /ug=Hs.79137 /len=1620	37736_at
MPHOSPH6 (M-phase phosphoprotein 6)	X98263	Hs.152720	NM_005792	16q24	Cluster Incl. X98263:H.sapiens mRNA for M-phase phosphoprotein, mpp6 /cds=(32,514) /gb=X98263 /gi=1770461 /ug=Hs.152720 /len=1079	31864_at
HCLS1 (hematopoietic cell-specific Lyn substrate 1)	X16663	Hs.14601	NM_005335	3q13	Cluster Incl. X16663:Human HS1 gene for heamatopoietic lineage cell specific protein /cds=(42,1502) /gb=X16663 /gi=32054 /ug=Hs.14601 /len=1950	31820_at

KIAA0451(KIAA0451 gene product)	AB007920			1	Cluster Incl. AB007920:Homo sapiens mRNA for KIAA0451 protein, complete cds /cds=(1482,2219) /gb=AB007920 /gi=3413863 /ug=Hs.18586 /len=6597	32206_at
CAT (catalase)	AL035079	Hs.76359	NM_001752	11p13	Cluster Incl. AL035079:dJ53C18.1 (Catalase) /cds=(74,1657) /gb=AL035079 /gi=4775614 /ug=Hs.76359 /len=2287	37009_at
OAZ1 (ornithine decarboxylase antizyme 1)	D78361	Hs.125078		19p13.3	D78361 /FEATURE= /DEFINITION=HUMODAZ Human mRNA for ornithine decarboxylase antizyme, ORF 1 and ORF 2	1315_at
VDAC3 (voltage-dependent anion channel 3)	AF038962	Hs.7381	NM_005662	8p11.2	Cluster Incl. AF038962:Homo sapiens voltage dependent anion channel protein mRNA, complete cds /cds=(99,950) /gb=AF038962 /gi=3329393 /ug=Hs.7381 /len=1384	36102_at
NCOA4 (nuclear receptor coactivator 4)	X77548	Hs.99908	NM_005437	10q11.2	Cluster Incl. X77548:H. sapiens cDNA for RFG /cds=(76,1920) /gb=X77548	39174_at

					/gi=469145 /ug=Hs.99908 /len=3418	
	D45288				Cluster Incl. D45288:HUMHG2121 Homo sapiens cDNA /gb=D45288 /gi=1136684 /ug=Hs.57079 /len=1479-	35310_at
RFP (ret finger protein)	J03407	Hs.142653	NM_006510	6p22	Cluster Incl. J03407:Human rfp transforming protein mRNA, complete cds /cds=(234,1775) /gb=J03407 /gi=337371 /ug=Hs.142653 /len=1782	40176_at
BRCA1 (breast cancer 1, early onset)	L78833	Hs.194143	NM_007295	17q21	L78833 /FEATURE=exon#24 /DEFINITION=HUMBRCA1 Human BRCA1, Rho7 and vatl genes, complete cds, and ipf35 gene, partial cds	604_at
BAZ1A (bromodomain adjacent to zinc finger domain, 1A)	AL050089	Hs.8858	NM_013448	14q12-q13	Cluster Incl. AL050089:Homo sapiens mRNA; cDNA DKFZp586E0518 (from clone DKFZp586E0518) /cds=(0,2435) /gb=AL050089 /gi=4884107 /ug=Hs.8858 /len=3215	37971_at

PPIE (peptidylprolyl isomerase E (cyclophilin E))	AF042386	Hs.33251	NM_006112	1p32	Cluster Incl. AF042386:Homo sapiens cyclophilin-33B (CYP-33) mRNA, complete cds /cds=(60,950) /gb=AF042386 /gi=2828150 /ug=Hs.33251 /len=1099	34366_g_at
CD163 (CD163 antigen)	Z22971	Hs.74076	NM_004244	12p13.3	Cluster Incl. Z22971:H.sapiens mRNA for M130 antigen extracellular variant /cds=(101,3550) /gb=Z22971 /gi=312147 /ug=Hs.166016 /len=3802	31438_s_at
RXRA (retinoid X receptor, alpha)	U66306	Hs.20084	NM_002957	9q34.3	Cluster Incl. U66306:Human retinoid X receptor alpha mRNA, 3 UTR, partial sequence /cds=UNKNOWN /gb=U66306 /gi=3411007 /ug=Hs.20084 /len=3772	32800_at
ADD2 (adducin 2 (beta))	U43959	Hs.247423	NM_017488	2p14-p13	Cluster Incl. U43959:Human beta 4 adducin mRNA, alternatively spliced partial cds /cds=(0,938) /gb=U43959 /gi=1172145 /ug=Hs.4852 /len=1284	36052_at

GOLTC1 (golgi transport complex 1 (90 kDa subunit))	AF058718	Hs.239631	NM_006348	7q31	Cluster Incl. AF058718:Homo sapiens putative 13 S Golgi transport complex 90kD subunit brain-specific isoform mRNA, complete cds /cds=(51,2570) /gb=AF058718 /gi=3808234 /ug=Hs.239631 /len=3105	34737_at
KIAA0089(KIAA0089 protein)	D42047	Hs.82432	-	3	Cluster Incl. D42047:Human mRNA for KIAA0089 gene, partial cds /cds=(0,1236) /gb=D42047 /gi=577306 /ug=Hs.82432 /len=4043	38394_at
TGFA (transforming growth factor, alpha)	X70340	Hs.170009	NM_003236	2p13	X70340 /FEATURE=cds /DEFINITION=HSTGFAA H.sapiens mRNA for transforming growth factor alpha /NOTE=replacement of probe set 383_at	160025_at
OGDH (oxoglutarate dehydrogenase (lipoamide))	D10523	Hs.168669	NM_002541	7p14-p13	Cluster Incl. D10523:Human mRNA for 2-oxoglutarate dehydrogenase, complete cds /cds=(57,3065) /gb=D10523 /gi=531240 /ug=Hs.168669 /len=4122	40470_at

IQGAP2 (IQ motif containing GTPase activating protein 2)	U51903	Hs.78993	NM_006633	5q	Cluster Incl. U51903:Human RasGAP-related protein (IQGAP2) mRNA, complete cds /cds=(222,4949) /gb=U51903 /gi=1262925 /ug=Hs.78993 /len=5767	37276_at
ASML3B(acid sphingomyelinase-like phosphodiesterase)	Y08134	Hs.123659	NM_014474	1	Cluster Incl. Y08134:H.sapiens mRNA for ASM-like phosphodiesterase 3b /cds=(121,1518) /gb=Y08134 /gi=1552274 /ug=Hs.123659 /len=1610	37779_at
GIPR (gastric inhibitory polypeptide receptor)	X81832	Hs.251412	NM_000164	19q13.3	Cluster Incl. X81832:H.sapiens mRNA for glucose-dependant insulinotropic polypeptide receptor gene /cds=(486,1961) /gb=X81832 /gi=1030050 /ug=Hs.142900 /len=2181	35590_s_at
ITSN(human intersectin-SH3 domain-containing protein SH3P17)	U61166	Hs.307177		21q22.11	U61166 /FEATURE= /DEFINITION=HSU61166 Human SH3 domain-containing protein SH3P17 mRNA, complete cds	488_at

KIAA1155(KIAA1155 protein)	AF090102	Hs.102657		2	Cluster Incl. AF090102:Homo sapiens clone IMAGE 21785 /cds=UNKNOWN /gb=AF090102 /gi=4063637 /ug=Hs.102657 /len=1712	39527_at
PLCE2 (phospholipase C, epsilon 2)	AB029015	Hs.54886		3p25.3-p25.1	Cluster Incl. AB029015:Homo sapiens mRNA for KIAA1092 protein, partial cds /cds=(0,3464) /gb=AB029015 /gi=5689520 /ug=Hs.54886 /len=4147	41796_at
DOC-1R(tumor suppressor deleted *in oral* cancer-related 1)	AF089814	Hs.25664	NM_005851	11	Cluster Incl. AF089814:Homo sapiens growth suppressor related (DOC-1R) mRNA, complete cds /cds=(103,483) /gb=AF089814 /gi=3661528 /ug=Hs.25664 /len=931	35151_at
IGL@ (immunoglobulin lambda locus)	M18645	Hs.181125		22q11.1-q11.2	Cluster Incl. M18645:Human Ig rearranged lambda-chain mRNA VJC-region subgroup lambda-IV *from* heterohybridoma H6-3C4 /cds=(30,731) /gb=M18645 /gi=186103 /ug=Hs.181125 /len=872	33274_f_at

	S71043				Cluster Incl. S71043:Ig alpha 2=immunoglobulin A heavy chain allotype 2 {constant region, germ line} [human, peripheral blood neutrophils, Genomic, 1799 nt] /cds=(0,1022) /gb=S71043 /gi=546798 /ug=Hs.32225 /len=1047	33501_r_at
UBE2V1 (ubiquitin-conjugating enzyme E2 variant 1)	U49278	Hs.75875	NM_003349	20q13.2	Cluster Incl. U49278:Homo sapiens UEV-1 (UBE2V) mRNA, partial cds /cds=(0,231) /gb=U49278 /gi=1709703 /ug=Hs.75875 /len=3335	36959_at
CBX7 (chromobox homolog 7)	AL031846			22q13.1	Cluster Incl. AL031846:dJ742C19.5 (novel Chromobox protein) /cds=(89,844) /gb=AL031846 /gi=4164368 /ug=Hs.7442 /len=3964	36894_at
	AL049675				Cluster Incl. AL049675:Human gene from PAC 886K2, chromosome 1 /cds=UNKNOWN /gb=AL049675 /gi=4678768 /ug=Hs.15535 /len=1074	32048_at

TPM1 (tropomyosin 1 (alpha))	M19267	Hs.77899	NM_000366	15q22.1	Cluster Incl. M19267:Human tropomyosin mRNA, complete cds /cds=(286,1140) /gb=M19267 /gi=339943 /ug=Hs.77899 /len=1633	36790_at
MYCBP (c-myc binding protein)	AB007191	Hs.78221	NM_012333	1p33-p32.2	Cluster Incl. AB007191:Homo sapiens mRNA for AMY-1, complete cds /cds=(38,349) /gb=AB007191 /gi=2443309 /ug=Hs.78221 /len=1492	37250_at
	AF052169				Cluster Incl. AF052169:Homo sapiens clone 24775 mRNA sequence /cds=UNKNOWN /gb=AF052169 /gi=3360480 /ug=Hs.109438 /len=1385	38972_at
TOP3A (topoisomerase (DNA) III alpha)	U43431	Hs.91175	NM_004618	17p12-17p11.2	U43431 /FEATURE= /DEFINITION=HSU43431 Human DNA topoisomerase III mRNA, complete cds	1028_at
PCAF (p300/CBP-associated factor)	U57317	Hs.199061	NM_003884	3p24	U57317 /FEATURE= /DEFINITION=HSU57317 Homo sapiens p300/CBP-associated factor (P/CAF)	1012_at

					mRNA, complete cds	
RAD54L (RAD54 (S.cerevisiae)-like)	X97795	Hs.66718	NM_003579	1p32	X97795 /FEATURE=cds /DEFINITION=HSRAD54 H.sapiens mRNA homologous to S. cerevisiae RAD54	966_at
CTSE (cathepsin E)	J05036	Hs.1355	NM_001910	1q31	J05036 /FEATURE=mRNA /DEFINITION=HUMCTSE Human cathepsin E mRNA, complete cds	271_s_at
IGLL3 (immunoglobulin lambda-like polypeptide 3)	AI932613	Hs.296552		22q11.23	Cluster Incl. AI932613:wo05c02.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-2454434 /clone_end=3 /gb=AI932613 /gi=5671350 /ug=Hs.62036 /len=570	41827_f_at
	S71043				Cluster Incl. S71043:Ig alpha 2=immunoglobulin A heavy chain allotype 2 {constant region, germ line} [human, peripheral blood neutrophils, Genomic, 1799 nt] /cds=(0,1022) /gb=S71043	33500_i_at

					/gi=546798 /ug=Hs.32225 /len=1047	
ZNF75 (zinc finger protein 75 (D8C6))	S67970	Hs.29159		xq26	Cluster Incl. S67970:ZNF75=KRAB zinc finger [human, lung fibroblast, mRNA, 1563 nt] /cds=UNKNOWN /gb=S67970 /gi=460902 /ug=Hs.29159 /len=1563	35222_at
PLSCR1 (phospholipid scramblase 1)	AB006746	Hs.198282	NM_021105	3q23	Cluster Incl. AB006746:Homo sapiens hMmTRA1b mRNA, complete cds /cds=(256,1212) /gb=AB006746 /gi=3510296 /ug=Hs.198282 /len=2077	32775_r_at
LOC51112(CGI-87 protein)	AI652660	Hs.5008	NM_016030		Cluster Incl. AI652660:wb30c10.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-2307186 /clone_end=3 /gb=AI652660 /gi=4736639 /ug=Hs.5008 /len=525	41590_at
	AF067420				Cluster Incl. AF067420:Homo sapiens SNC73 protein (SNC73) mRNA, complete cds /cds=(395,1549) /gb=AF067420	33499_s_at

					/gi=3201899 /ug=Hs.32225 /len=1594	
CSTF3 (cleavage stimulation factor, 3' pre-RNA, subunit 3, 77kD)	U15782	Hs.180034	NM_001326	11	Cluster Incl. U15782:Human cleavage stimulation factor 77kDa subunit mRNA, complete cds /cds=(131,2284) /gb=U15782 /gi=632497 /ug=Hs.180034 /len=2766	41183_at
CAPZA1 (capping protein (actin filament) muscle Z-line, alpha 1)	U56637	Hs.184270	NM_006135	1p36.13-q23.3	Cluster Incl. U56637:Human capping protein alpha subunit isoform 1 mRNA, complete cds /cds=(0,860) /gb=U56637 /gi=1336098 /ug=Hs.184270 /len=2366	40910_at
NEDD4L (reserved)	AB007899	Hs.12017	NM_015277	18q21	Cluster Incl. AB007899:Homo sapiens KIAA0439 mRNA, partial cds /cds=(0,2989) /gb=AB007899 /gi=2662158 /ug=Hs.12017 /len=4879	39356_at
CCNB1 (cyclin B1)	M25753:	Hs.23960	NM_031966	5q12	Cluster Incl. M25753:Human cyclin B mRNA, 3 end /cds=UNKNOWN /gb=M25753 /gi=181243 /ug=Hs.23960	34736_at

					/len=1452	
ITSN1 (intersectin 1 (SH3 domain protein))	AF064243	Hs.66392	NM_003024	21q22.1-q22.2	Cluster Incl. AF064243:Homo sapiens intersectin short form mRNA, complete cds /cds=(106,3768) /gb=AF064243 /gi=3859852 /ug=Hs.66392 /len=5272	35776_at
BCS1L (BCS1 (yeast homolog)-like)	AF038195	Hs.150922	NM_004328	2q33	Cluster Incl. AF038195:Homo sapiens clone 23661 unknown protein mRNA, complete cds /cds=(75,1334) /gb=AF038195 /gi=2795915 /ug=Hs.150922 /len=1391	31842_at
KIAA0229(KIAA0229 protein)	D86982	Hs.20060		6	Cluster Incl. D86982:Human mRNA for KIAA0229 gene, partial cds /cds=(0,3543) /gb=D86982 /gi=1504037 /ug=Hs.20060 /len=6335	40971_at
RNF24 (ring finger protein 24)	AL031670	Hs.30524	NM_007219	20p13-p12.1	Cluster Incl. AL031670:dJ681N20.2 (ferritin, light polypeptide-like 1) /cds=(200,727) /gb=AL031670	35083_at

					/gi=4469083 /ug=Hs.111334 /len=978	
SLC29A2 (solute carrier family 29 (nucleoside transporters), member 2)	AF034102	Hs.32951	NM_001532	11q13	Cluster Incl. AF034102:Homo sapiens NBMPR-insensitive nucleoside transporter ei (ENT2) mRNA, complete cds /cds=(237,1607) /gb=AF034102 /gi=2811136 /ug=Hs.32951 /len=2522	39661_s_at
KIAA0436(putative L-type neutral amino acid transporter)]	AB007896			2	Cluster Incl. AB007896:Homo sapiens KIAA0436 mRNA, partial cds /cds=(0,2069) /gb=AB007896 /gi=2662152 /ug=Hs.110 /len=4661	38984_at
ACTB (actin, beta)	X00351	Hs.288061	NM_001101	7p15-p12	Homo sapiens /REF=X00351 /DEF=Human mRNA for beta-actin /LEN=1761 (_5, _M, _3 represent transcript regions 5 prime, Middle, and 3 prime respectively)	AFFX-HSAC07
CSPG2 (chondroitin sulfate proteoglycan 2 (versican))	D32039	Hs.81800	NM_004385	5q14.3	Cluster Incl. D32039:Human pgH3 mRNA for proteoglycan PG-M(V3), complete cds /cds=(105,2072) /gb=D32039 /gi=1008912	31682_s_at

					/ug=Hs.234753 /len=2087	
SLC29A1 (solute carrier family 29 (nucleoside transporters), member 1	U81375	Hs.25450	NM_004955	6p21.1-p21.2	Cluster Incl. U81375:Human placental equilibrative nucleoside transporter 1 (hENT1) mRNA, complete cds /cds=(178,1548) /gb=U81375 /gi=1845344 /ug=Hs.25450 /len=2162	33901_at
PVR (poliovirus receptor)	X64116	Hs.321018	NM_006505	19q13.2	Cluster Incl. X64116:H.sapiens PVR gene for poliovirus receptor (exon 1) /cds=(205,1299) /gb=X64116 /gi=35809 /ug=Hs.171844 /len=1300	32699_s_at
KNSL3 (kinesin-like 3)	AB012722	Hs.198256	NM_030615	6q27	Cluster Incl. AB012722:Homo sapiens gene for kinesin-like protein, complete cds /cds=(94,1248) /gb=AB012722 /gi=4115550 /ug=Hs.198256 /len=1342	31978_at
13CDNA73(putative gene product)	U50534	Hs.181304	NM_023037	13	U50534 /FEATURE= /DEFINITION=HSU50534 Human BRCA2 region, mRNA sequence CG003	1529_at

ITGAV (integrin, alpha V (vitronectin receptor, alpha polypeptide, antigen CD51))	U07375	Hs.295726	NM_002210	2q31-q32	U07375 /FEATURE=cds /DEFINITION=HSU07375 Human integrin alpha v gene, promoter region and partial cds	2032_s_at
DPYD (dihydropyrimidine dehydrogenase)	U20938	Hs.1602	NM_000110	1p22	Cluster Incl. U20938:Human lymphocyte dihydropyrimidine dehydrogenase mRNA, complete cds /cds=(101,3178) /gb=U20938 /gi=1926407 /ug=Hs.1602 /len=4409	38220_at
KIAA0152(KIAA0152 gene product)	D63486	- Hs.181418	NM_014730	12	Cluster Incl. D63486:Human mRNA for KIAA0152 gene, complete cds /cds=(128,1006) /gb=D63486 /gi=1469885 /ug=Hs.181418 /len=6322	41728_at

Table 5:

UCL/HGNC/HUGO Human Gene Nomenclature Database Symbol	GenBank Accession No.	UniGene Cluster	RefSeq	Chromosomal Location	Description Unigene Build #95	Gene Name
GATA2 (GATA-binding protein 2)	M77810	Hs.334695	NM_002050	3q21	M77810 /FEATURE= /DEFINITION=HUMGATA2A Human transcription factor GATA-2 (GATA-2) mRNA, complete cds	1072_g_at
PRG2 (proteoglycan 2, bone marrow (natural killer cell activator, eosinophil granule major basic	Z26248	Hs.99962	NM_002728	11q12	Cluster Incl. Z26248:H.sapiens mRNA for eosinophil granule major basic protein /cds=(857,1525) /gb=Z26248 /gi=940510 /ug=Hs.99962 /len=1637	39179_at
FAT2 (FAT tumor suppressor (Drosophila) homolog 2)	AB011535	Hs.158159	NM_001447	5q32-q33	Cluster Incl. AB011535:Homo sapiens mRNA for MEGF1, partial cds /cds=(0,1721) /gb=AB011535 /gi=3449295 /ug=Hs.158159 /len=3193	38202_at

PROC (protein C (inactivator of coagulation factors Va and VIIIa))	X02750	Hs.2351	NM_000312	2q13-q14	Cluster Incl. X02750:Human liver mRNA for protein C /cds=(97,1482) /gb=X02750 /gi=35689 /ug=Hs.2351 /len=1843	39255_at
	AC005764				Cluster Incl. AC005764:Homo sapiens chromosome 19, cosmid R31343 /cds=(0,1262) /gb=AC005764 /gi=3694626 /ug=Hs.126496 /len=1263	35512_at
CAMK2B (calcium/calmodulin-dependent protein kinase (CaM kinase) II beta)	AF112471	Hs.4884	NM_001220	7p14.3-p14.1	Cluster Incl. AF112471:Homo sapiens calcium/calmodulin-dependent protein kinase II beta subunit mRNA, alternatively spliced, complete cds /cds=(46,1599) /gb=AF112471 /gi=4139267 /ug=Hs.4884 /len=1750	34847_s_at
BRF2 (butyrate response factor 2 (EGF-response factor 2))	X78992	Hs.78909	NM_006887	2p22.3-2p21	Cluster Incl. X78992:H.sapiens ERF-2 mRNA /cds=(66,1544) /gb=X78992 /gi=509777 /ug=Hs.78909 /len=1629	32588_s_at
RPL7 (ribosomal protein L7)	X57958	Hs.153	NM_000971	8q	Cluster Incl. X57958:H.sapiens mRNA for ribosomal protein L7 /cds=(22,783)	36333_at

					/gb=X57958 /gi=35904 /ug=Hs.153 /len=847	
GADD45A (growth arrest and DNA-damage-inducible, *alpha*)	M60974	Hs.80409	NM_001924	1p31.2-p31.1	M60974 /FEATURE= /DEFINITION=HUMGADD45 Human growth arrest and DNA-damage-inducible protein (gadd45) mRNA, complete cds	1911_s_at
NUDEL(*nuclear distribution gene E-like*)]	AF038203	Hs.3850	NM_030808	17	Cluster Incl. AF038203:Homo sapiens clone 23596 mRNA sequence /cds=UNKNOWN /gb=AF038203 /gi=2795924 /ug=Hs.3850 /len=1473	34413_at
MAD (MAX dimerization protein)	L06895	Hs.109012	NM_002357	2p13-p12	L06895 /FEATURE= /DEFINITION=HUMMAD Homo sapiens antagonizer of myc transcriptional activity (Mad) mRNA, complete cds	1774_at
LRPAP1 (*low density lipoprotein-related protein-associated protein 1 (alpha-2-macroglobulin*	M63959	Hs.75140	NM_002337	4p16.3	Cluster Incl. M63959:Human alpha-2-macroglobulin receptor-associated protein mRNA, complete cds /cds=(13,1086) /gb=M63959 /gi=177873 /ug=Hs.75140	36194_at

					/len=1493	
RPL22 (ribosomal protein L22)	AI526079	Hs.326249	NM_000983	3q26	Cluster Incl. AI526079:DU3.2-7.G09 Homo sapiens cDNA, 3 end /clone_end=3 /gb=AI526079 /gi=4440197 /ug=Hs.234060 /len=801	33451_s_at
CLC (Charot-Leyden crystal protein)	L01664	Hs.132004	NM_013246	11q13.3	Cluster Incl. L01664:Human eosinophil Charcot-Leyden crystal (CLC) protein (lysophospholipase) mRNA, complete cds /cds=(33,461) /gb=L01664 /gi=187273 /ug=Hs.889 /len=586	36809_at
GFI1 (growth factor independent 1)	U67369	Hs.73172	NM_005263	1p22	Cluster Incl. U67369:Human growth factor independence-1 (Gfi-1) mRNA, complete cds /cds=(267,1535) /gb=U67369 /gi=1698691 /ug=Hs.73172 /len=2799	33977_at
MYL6 (myosin, light polypeptide 6, alkali, smooth muscle and non-muscle)	M22919	Hs.77385	NM_021019	12	Cluster Incl. M22919:Human nonmuscle/smooth muscle alkali myosin light chain gene, complete cds	33994_g_at

					/cds=(42,353) /gb=M22919 /gi=189016 /ug=Hs.77385 /len=1259	
[EBBP(tripartite motif protein 16)]	AF096870	Hs.241305	NM_006470	17	Cluster Incl. AF096870:Homo sapiens estrogen-responsive B box protein (EBBP) mRNA, complete cds /cds=(227,1921) /gb=AF096870 /gi=3916726 /ug=Hs.194540 /len=2568	38881_i_at
RNASE2 (ribonuclease, RNase A family, 2 (liver, eosinophil-derived neurotoxin))	X55988	Hs.728	NM_002934	14q24-q31	Cluster Incl. X55988:Human EDN mRNA for eosinophil derived neurotoxin /cds=(71,556) /gb=X55988 /gi=31088 /ug=Hs.728 /len=735	36766_at
RBMS3 (RNA binding motif, single stranded interacting protein)	AA523313	Hs.158446	NM_014483	3p24-p23	Cluster Incl. AA523313:ni41h09.s1 Homo sapiens cDNA, 3 end /clone=IMAGE-979457 /clone_end=3 /gb=AA523313 /gi=2264025 /ug=Hs.158446 /len=581	33583_r_at
ATBF1 (AT-binding transcription factor 1)	L32832	Hs.101842	NM_006885	16q22.3-q23.1	Cluster Incl. L32832:Homo sapiens zinc finger homeodomain protein (ATBF1-A) mRNA, complete cds /cds=(673,11784)	37114_at

					/gb=L32832 /gi=976346 /ug=Hs.101842 /len=11893	
ALDOA (aldolase A, fructose-bisphosphate)	X05236	Hs.273415	NM_000034	16q22-q24	Cluster Incl. X05236:Human fibroblast mRNA for aldolase A /cds=(146,1240) /gb=X05236 /gi=28596 /ug=Hs.183760 /len=1440	32336_at
DSIPI (delta sleep inducing peptide, immunoreactor)	AI635895	Hs.75450	NM_004089	xp21.1-q25	Cluster Incl. AI635895:tz82a07.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-2295060 /clone_end=3 /gb=AI635895 /gi=4687225 /ug=Hs.75450 /len=1082	36629_at
BPI (bactericidal/permeability-increasing protein)	J04739	Hs.89535	NM_001725	20q11.23-q12	Cluster Incl. J04739:Human bactericidal permeability increasing protein (BPI) mRNA, complete cds /cds=(30,1493) /gb=J04739 /gi=179528 /ug=Hs.89535 /len=1813	37054_at
JUND (jun D proto-oncogene)	X56681	Hs.2780	NM_005354	19p13.2	Cluster Incl. X56681:Human junD mRNA /cds=(174,1217) /gb=X56681 /gi=34018	41483_s_at

					/ug=Hs.2780 /len=1891	
MYB (v-myb avian myeloblastosis viral oncogene homolog)	M13666	Hs.1334	NM_005375	6q22-q23	Cluster Incl. M13666:Human c-myb mRNA, 3 end /cds=(0,833) /gb=M13666 /gi=180657 /ug=Hs.1334 /len=1035	41854_at
EPX (eosinophil peroxidase)	X14346	Hs.46295	NM_000502	17q23.1	Cluster Incl. X14346:Human mRNA for eosinophil peroxidase /cds=(0,2108) /gb=X14346 /gi=31182 /ug=Hs.46295 /len=2558	34587_at
PGD (phosphogluconate dehydrogenase)	U30255	Hs.75888	NM_002631	1p36.3-p36.13	Cluster Incl. U30255:Human phosphogluconate dehydrogenase (hPGDH) gene, complete cds /cds=(6,1457) /gb=U30255 /gi=984324 /ug=Hs.75888 /len=1536	36963_at
ATF2 (activating transcription factor 2)	U16028	Hs.198166	NM_001880	2q32	U16028 /FEATURE= /DEFINITION=HSU16028 Human CRE-BP1 transcription factor mRNA, complete cds	1994_at

TAGLN2 (transgelin 2)	D21261	Hs.75725	NM_003564	1q21-q25	Homo sapiens /REF=D21261 /DEF=Cluster Incl. :Human mRNA for KIAA0120 gene, complete cds /cds=(73,672) /gb= /gi=434762 /ug=Hs.75725 /len=1360 /LEN=1594	36678_at
GP9 (glycoprotein IX (platelet))	X52997	Hs.1144	NM_000174	3q21	Cluster Incl. X52997:Human mRNA for platelet glycoprotein IX /cds=(222,755) /gb=X52997 /gi=2160045 /ug=Hs.1144 /len=888	35101_at
GAPD (glyceraldehyde-3-phosphate dehydrogenase)	M33197	Hs.169476	NM_002046	12p13	Homo sapiens /REF=M33197 /DEF=Human glyceraldehyde-3-phosphate dehydrogenase (GAPDH) mRNA, complete cds /LEN=1268 (_5, _M, _3 represent transcript regions 5 prime, Middle, and 3 prime respectively)	AFFX-HUMGA
LMOD1 (leiomodin 1 (smooth muscle))	X54162	Hs.79386	NM_012134	1q32	Cluster Incl. X54162:Human mRNA for a 64 Kd autoantigen expressed in thyroid and extra-ocular muscle /cds=(212,1930) /gb=X54162 /gi=28968 /ug=Hs.79386	37765_at

					/len=3849	
JUND (jun D proto-oncogene)	X56681	Hs.2780	NM_005354	19p13.2	X56681 /FEATURE=mRNA /DEFINITION=HSJUNDR Human junD mRNA	1612_s_at
CCNI (cyclin I)	D50310	Hs.79933	NM_006835	4	D50310 /FEATURE= /DEFINITION=HUMCYI Human mRNA for cyclin I, complete cds	1836_at
RPS4X (*ribosomal protein S4, X-linked*)	M58458	Hs.108124	NM_001007	xq13.1	Cluster Incl. M58458:Human ribosomal protein S4 (RPS4X) isoform mRNA, complete cds /cds=(35,826) /gb=M58458 /gi=337509 /ug=Hs.75344 /len=888	34643_at
RPS6 (ribosomal protein S6)	X67309	Hs.241507	NM_001010	9p21	Cluster Incl. X67309:H.sapiens gene for ribosomal protein S6 /cds=(42,791) /gb=X67309 /gi=36147 /ug=Hs.120856 /len=829	35125_at

MAD (MAX dimerization protein)	L06895	Hs.109012	NM_002357	2p13-p12	Cluster Incl. L06895:Homo sapiens antagonizer of myc transcriptional activity (Mad) mRNA, complete cds /cds=(147,812) /gb=L06895 /gi=187288 /ug=Hs.239794 /len=1002	34543_at
H2AFY (H2A histone family, member Y)	AF054174	Hs.75258	NM_004893	5q31.3-q32	Cluster Incl. AF054174:Homo sapiens histone macroH2A1.2 mRNA, complete cds /cds=(173,1288) /gb=AF054174 /gi=3341991 /ug=Hs.75258 /len=1881	36576_at
SAT (spermidine/spermine N1-acetyltransferase)		Hs.28491	NM_002970	xp22.1	Spermidine/Spermine N1-Acetyltransferase, Alt. Splice 2	1173_g_at
IGHG3 (immunoglobulin heavy constant gamma 3 (G3m marker))	AI147237	Hs.300697		14q32.33	Cluster Incl. AI147237:qb36f02.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-1698363 /clone_end=3 /gb=AI147237 /gi=3674919 /ug=Hs.210732 /len=474	34104_i_at
KIAA0225(KIAA0225 protein	D86978	Hs.84790		7	Cluster Incl. D86978:Human mRNA for KIAA0225 gene, partial cds /cds=(0,6043) /gb=D86978 /gi=1504029 /ug=Hs.84790	38728_at

					/len=6237	
HMG4 (high-mobility group (nonhistone chromosomal) protein 4)	AL034450	Hs.19114	NM_005342	Xq28	Cluster Incl. AL034450:Human DNA sequence from clone 115K14 on chromosome Xq22.3-23 Contains high mobility group protein 2a, ESTs, STS /cds=(0,605) /gb=AL034450 /gi=4210359 /ug=Hs.194749 /len=730	31588_at
MSF (MLL septin-like fusion (NOTE: non-standard symbol and name))	AB023208	Hs.181002	NM_006640	17q25	Cluster Incl. AB023208:Homo sapiens mRNA for KIAA0991 protein, complete cds /cds=(732,2000) /gb=AB023208 /gi=4589625 /ug=Hs.181002 /len=3938	41220_at
CRA(cisplatin resistance associated)	U78556	Hs.166066	NM_006697	1	U78556 /FEATURE= /DEFINITION=HSU78556 Human cisplatin resistance associated alpha protein (hCRA alpha) mRNA, complete cds	1229_at
	AF052169				Cluster Incl. AF052169:Homo sapiens clone 24775 mRNA sequence /cds=UNKNOWN /gb=AF052169	38972_at

					/gi=3360480 /ug=Hs.109438 /len=1385	
CD164 (CD164 antigen, sialomucin)	D14043	Hs.43910	NM_006016	6q21	Cluster Incl. D14043:Human mRNA for MGC-24, complete cds /cds=(79,648) /gb=D14043 /gi=219924 /ug=Hs.43910 /len=2417	34819_at
PLXNB2 (plexin B2)	AB002313	Hs.3989		22q13.33	Cluster Incl. AB002313:Human mRNA for KIAA0315 gene, partial cds /cds=(0,5526) /gb=AB002313 /gi=2280475 /ug=Hs.3989 /len=6252	34780_at
KIAA1042(KIAA1042 protein)	AB028965	Hs.6705	NM_014965	3	Cluster Incl. AB028965:Homo sapiens mRNA for KIAA1042 protein, complete cds /cds=(216,3077) /gb=AB028965 /gi=5689420 /ug=Hs.6705 /len=5109	35789_at
CSPG2 (chondroitin sulfate proteoglycan 2 (versican))	X15998	Hs.81800	NM_004385	5q14.3	Cluster Incl. X15998:H.sapiens mRNA for the chondroitin sulphate proteoglycan versican, V1 splice-variant; precursor peptide /cds=(266,7495) /gb=X15998	38111_at

					/gi=37662 /ug=Hs.81800 /len=8224	
DCK (deoxycytidine kinase)	M60527	Hs.709	NM_000788	4q13.3-q21.1	M60527 /FEATURE=mRNA /DEFINITION=HUMDCKATPB Human deoxycytidine kinase mRNA, complete cds	886_at
TGFBI (transforming growth factor, beta-induced, 68kD)	M77349	Hs.118787	NM_000358	5q31	M77349 /FEATURE= /DEFINITION=HUMTGFBIG Human transforming growth factor-beta induced gene product (BIGH3) mRNA, complete cds	1385_at
SIAT9 (sialyltransferase 9 (CMP-NeuAc:lactosylceramide alpha-2,3-sialyltransferase; GM3	AB018356	Hs.225939	NM_003896	2p24.3-p24.1	Cluster Incl. AB018356:Homo sapiens mRNA for GM3 synthase, complete cds /cds=(277,1365) /gb=AB018356 /gi=3779138 /ug=Hs.225939 /len=2359	34256_at
H2BFL (H2B histone family, member L)	AI688098	Hs.239884	NM_003526	6p21.3	Cluster Incl. AI688098:wc92f08.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-2326119 /clone_end=3 /gb=AI688098	33458_r_at

					/gi=4899392 /ug=Hs.239884 /len=576	
SLU7(step II splicing factor SLU7)	AI660656	Hs.76325	NM_006425	5	Cluster Incl. AI660656:wf23c07.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-2351436 /clone_end=3 /gb=AI660656 /gi=4764239 /ug=Hs.76325 /len=522	37006_at
KIAA0143(KIAA0143 protein)	D63477	Hs.84087		8	Cluster Incl. D63477:Human mRNA for KIAA0143 gene, partial cds /cds=(0,2658) /gb=D63477 /gi=1469867 /ug=Hs.84087 /len=5286	38472_at
CALM3 (calmodulin 3 (phosphorylase kinase, delta))	J04046	Hs.334330	NM_005184	19q13.2-q13.3	J04046 /FEATURE=mRNA /DEFINITION=HUMCAMA Human calmodulin mRNA, complete cds	1158_s_at
IGKC (immunoglobulin kappa constant)	M63438	Hs.156110	-	2p12	Cluster Incl. M63438:Human Ig rearranged gamma chain mRNA, V-J-C region and complete cds /cds=(0,1049) /gb=M63438 /gi=184847 /ug=Hs.156110 /len=1244	38194_s_at

IRF4 (interferon regulatory factor 4)	U52682	Hs.82132	NM_002460	6p25-p23	Cluster Incl. U52682:Human lymphocyte specific interferon regulatory factor/interferon regulatory factor 4 (LSIRF/IRF4) mRNA, complete cds /cds=(125,1477) /gb=U52682 /gi=1378108 /ug=Hs.82132 /len=5320	37625_at
HLA-DMB (major histocompatibility complex, class II, DM beta)	U15085	Hs.1162	NM_002118	6p21.3	Cluster Incl. U15085:Human HLA-DMB mRNA, complete cds /cds=(233,1024) /gb=U15085 /gi=557701 /ug=Hs.1162 /len=1362	41609_at
IGHM (immunoglobulin heavy constant mu)	X67301	Hs.302063		14q32.33	Cluster Incl. X67301:H.sapiens mRNA for IgM heavy chain constant region (Ab63) /cds=(0,1361) /gb=X67301 /gi=38407 /ug=Hs.179543 /len=1453	41164_at
MTMR1 (myotubularin related protein 1)	AJ224979	Hs.23200		Xq28	Cluster Incl. AJ224979:Homo sapiens mRNA for MTMR1 protein /cds=(0,1990) /gb=AJ224979 /gi=4128155 /ug=Hs.23200 /len=2582	34654_at

LILRB4 (leukocyte immunoglobulin-like receptor, subfamily B (with TM and ITIM domains),	AF072099	Hs.67846	NM_006847	19q13.4	Cluster Incl. AF072099:Homo sapiens immunoglobulin-like transcript 3 protein variant 1 gene, complete cds /cds=(0,1346) /gb=AF072099 /gi=3776463 /ug=Hs.67846 /len=1705	36753_at
SLC31A1 (solute carrier family 31 (copper transporters), member 1)	U83460	Hs.73614	NM_001859	9q31-q32	Cluster Incl. U83460:Human high-affinity copper uptake protein (hCTR1) mRNA, complete cds /cds=(152,724) /gb=U83460 /gi=2315986 /ug=Hs.73614 /len=1804	40364_at
CST3 (cystatin C (amyloid angiopathy and cerebral hemorrhage))	AI362017	Hs.135084	NM_000099	20p11.2	Cluster Incl. AI362017:qy39a10.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-2014362 /clone_end=3 /gb=AI362017 /gi=4113638 /ug=Hs.135084 /len=778	39689_at
KIAA0261(KIAA0261 protein)	D87450	Hs.154978		10	Cluster Incl. D87450:Human mRNA for KIAA0261 gene, partial cds /cds=(0,3865) /gb=D87450 /gi=1665788 /ug=Hs.154978 /len=6155	40086_at

ASB1 (ankyrin repeat and SOCS box-containing 1)	AF055024	Hs.153489	NM_016114	2q37	Cluster Incl. AF055024:Homo sapiens clone 24763 mRNA sequence /cds=UNKNOWN /gb=AF055024 /gi=3005752 /ug=Hs.153489 /len=1830	31875_at
HNRPA3 (heterogeneous nuclear ribonucleoprotein A3)	S63912	Hs.249247	NM_005758	10	Cluster Incl. S63912:D10S102=FBRNP [human, fetal brain, mRNA, 3043 nt] /cds=(30,839) /gb=S63912 /gi=399757 /ug=Hs.234462 /len=3043	33817_at
HLA-DQB1 (major histocompatibility complex, class II, DQ beta 1)	M81141	Hs.73931	NM_002123	6p21.3	Cluster Incl. M81141:Human MHC class II HLA-DQ-beta mRNA (DR7 DQw2), complete cds /cds=(35,820) /gb=M81141 /gi=188202 /ug=Hs.73933 /len=1171	36773_f_at
MD-1(MD-1, RP105-associated)	AB020499	Hs.184018	NM_004271	6	Cluster Incl. AB020499:Homo sapiens BCG-regulated mRNA for MD-1 homologue, complete cds /cds=(131,358) /gb=AB020499 /gi=4586549 /ug=Hs.184018 /len=713	35869_at

POLR2B (polymerase (RNA) II (DNA directed) polypeptide B (140kD))	X63563	Hs.296014	NM_000938	4q12	Cluster Incl. X63563:H.sapiens mRNA for RNA polymerase II 140 kDa subunit /cds=(43,3567) /gb=X63563 /gi=36121 /ug=Hs.148027 /len=3748	39746_at
	M13560				Cluster Incl. M13560:Human la-associated invariant gamma-chain gene /cds=(795,1493) /gb=M13560 /gi=184518 /ug=Hs.84298 /len=2080	35016_at
ITGB1 (integrin, beta 1 (fibronectin receptor, beta polypeptide, antigen CD29 includes MDF2, MSK12))	X07979	Hs.287797		10p11.2	Cluster Incl. X07979:Human mRNA for integrin beta 1 subunit /cds=(103,2499) /gb=X07979 /gi=31441 /ug=Hs.202661 /len=3614	32808_at
TRRAP (transformation/transcription domain-associated protein)	AF110377	Hs.203952	NM_003496	7q21.2-q22.1	Cluster Incl. AF110377:Homo sapiens PCAF-associated factor 400 (PAF400) mRNA, complete cds /cds=(129,11708) /gb=AF110377 /gi=4151928 /ug=Hs.203952 /len=12603	33810_at

SLC7A6 (solute carrier family 7 (cationic amino acid transporter, y+ system), member 6	D87432	Hs.10315	NM_003983	16q22.1-q22.3	Cluster Incl. D87432:Human mRNA for KIAA0245 gene, complete cds /cds=(261,1808) /gb=D87432 /gi=1665758 /ug=Hs.10315 /len=6296	39533_at
IGHM (immunoglobulin heavy constant mu)	X67301	Hs.302063		14q32.33	Cluster Incl. X67301:H.sapiens mRNA for IgM heavy chain constant region (Ab63) /cds=(0,1361) /gb=X67301 /gi=38407 /ug=Hs.179543 /len=1453	41165_g_at
CPVL (carboxypeptidase, vitellogenic-like)	AC005162	Hs.95594	NM_031311	7p15-p14	Cluster Incl. AC005162:Homo sapiens BAC clone RG113D17 from 7p14-p15 /cds=(0,887) /gb=AC005162 /gi=3242752 /ug=Hs.95594 /len=888	38323_at
IGL@ (immunoglobulin lambda locus)	M18645	Hs.181125		22q11.1-q11.2	Cluster Incl. M18645:Human Ig rearranged lambda-chain mRNA VJC-region subgroup lambda-IV from heterohybridoma H6-3C4 /cds=(30,731) /gb=M18645 /gi=186103 /ug=Hs.181125 /len=872	33274_f_at

DKFZP564B0769(DKFZP564B0769 protein)	AL080186	Hs.18368		6	Cluster Incl. AL080186:Homo sapiens mRNA; cDNA DKFZp564B0769 (from clone DKFZp564B0769) /cds=(0,900) /gb=AL080186 /gi=5262664 /ug=Hs.18368 /len=2776	41784_at
PRKWNK1 (protein kinase, lysine deficient 1)	U00946	Hs.184592	NM_018979	12p13.3	Cluster Incl. U00946:Human clone A9A2BRB5 (CAC)n/(GTG)n repeat-containing mRNA /cds=UNKNOWN /gb=U00946 /gi=405048 /ug=Hs.184592 /len=1971	32185_at
CD14 (CD14 antigen)	X06882	Hs.75627	NM_000591	5q31.1	Cluster Incl. X06882:Human gene for CD14 differentiation antigen /cds=(105,1232) /gb=X06882 /gi=29736 /ug=Hs.75627 /len=1356	36661_s_at
ALDH9A1 (aldehyde dehydrogenase 9 family, member A1)	U34252	Hs.2533	NM_000696	1q22-q23	Cluster Incl. U34252:Human gamma-aminobutyraldehyde dehydrogenase mRNA, complete cds /cds=(377,1858) /gb=U34252 /gi=1049218 /ug=Hs.2533	33899_at

					/len=2688	
MRPS18-2(mitochondrial ribosomal protein S18-2)	AL050361	Hs.274417	NM_014046	6	Cluster Incl. AL050361:Homo sapiens mRNA; cDNA DKFZp564H0223 (from clone DKFZp564H0223) /cds=UNKNOWN /gb=AL050361 /gi=4914594 /ug=Hs.190161 /len=1608	32221_at
KSR (kinase suppressor of ras)	U43586	Hs.152094		17q11.2	U43586 /FEATURE= /DEFINITION=HSU43586 Human kinase suppressor of ras-1 (KSR1) mRNA, partial cds	1716_at
					Cluster Incl. AB002448:Homo sapiens mRNA from chromosome 5q21-22, clone-357Ex /cds=UNKNOWN /gb=AB002448 /gi=2943811 /ug=Hs.26968 /len=1270	36260_at
KIAA0471(KIAA0471 gene product)	AB007940			1	Cluster Incl. AB007940:Homo sapiens mRNA for KIAA0471 protein, complete cds /cds=(412,1524) /gb=AB007940	34445_at

					/gi=3413903 /ug=Hs.107325 /len=6834	
TRB@ (T cell receptor beta locus)	M12886	Hs.303157		7q35	M12886 /FEATURE= /DEFINITION=HUMTCBYY Human T-cell receptor active beta-chain mRNA, complete cds	1105_s_at
					Cluster Incl. X92997:H.sapiens mRNA for IgG lambda light chain V-J-C region (clone Tgl4) /cds=(0,321) /gb=X92997 /gi=1070337 /ug=Hs.129722 /len=322	35530_f_at
IGLL3 (immunoglobulin lambda-like polypeptide 3)	AI932613	Hs.296552		22q11.23	Cluster Incl. AI932613:wo05c02.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-2454434 /clone_end=3 /gb=AI932613 /gi=5671350 /ug=Hs.62036 /len=570	41827_f_at
SFTPC (surfactant, pulmonary-associated protein C)	J03553	Hs.1074	NM_003018	8p21	Cluster Incl. J03553:Human pulmonary surfactant protein (SP5) mRNA, complete cds /cds=(146,739) /gb=J03553 /gi=338306 /ug=Hs.1074 /len=963	38691_s_at

SMAP(thyroid hormone receptor coactivating protein)	AW020536	Hs.5464	NM_006696	5	Cluster Incl. AW020536:df11b12.y1 Homo sapiens cDNA, 5 end /clone=IMAGE-2482918 /clone_end=5 /gb=AW020536 /gi=5874066 /ug=Hs.169344 /len=514	32653_at
KIAA0982(KIAA0982 protein)	AB023199	Hs.27207	NM_014023	10	Cluster Incl. AB023199:Homo sapiens mRNA for KIAA0982 protein, complete cds /cds=(144,1628) /gb=AB023199 /gi=4589607 /ug=Hs.27207 /len=4586	35199_at
AQP3 (aquaporin 3)	N74607	Hs.234642	NM_004925	9p13	Cluster Incl. N74607:za55a01.s1 Homo sapiens cDNA, 3 end /clone=IMAGE-296424 /clone_end=3 /gb=N74607 /gi=1231892 /ug=Hs.234642 /len=487	39248_at
					Cluster Incl. AL050166:Homo sapiens mRNA; cDNA DKFZp586D1122 (from clone DKFZp586D1122) /cds=UNKNOWN /gb=AL050166 /gi=4884381 /ug=Hs.26295 /len=2654	39582_at

NAGA (N-acetylgalactosaminidase, alpha-)	Z99716	Hs.75372	NM_000262	22q11	Cluster Incl. Z99716:bK250D10.5 (alpha-N-acetylgalactosaminidase) /cds=(472,1707) /gb=Z99716 /gi=4456457 /ug=Hs.75372 /len=3606	36607_at
CLTC (clathrin, heavy polypeptide (Hc))	D21260	Hs.178710	NM_004859	17q11-qter	Cluster Incl. D21260:Human mRNA for KIAA0034 gene, complete cds /cds=(172,5199) /gb=D21260 /gi=434760 /ug=Hs.178710 /len=6111	41159_at
MGEA5 (meningioma expressed antigen 5 (hyaluronidase))	AB014579	Hs.5734	NM_012215	10q24.1-q24.3	Cluster Incl. AB014579:Homo sapiens mRNA for KIAA0679 protein, partial cds /cds=(0,2303) /gb=AB014579 /gi=3327171 /ug=Hs.5734 /len=4303	35317_at
ANXA5 (annexin A5)	U05770	Hs.300711	NM_001154	4q28-q32	Cluster Incl. U05770:Human annexin V (ANX5) gene /cds=(164,1126) /gb=U05770 /gi=2182176 /ug=Hs.79274 /len=1597	37747_at
HLA-DRA (major histocompatibility complex, class II, DR alpha)	J00194	Hs.76807	NM_019111	6p21.3	Cluster Incl. J00194:human hla-dr antigen alpha-chain mrna & ivs fragments	37039_at

class II, DR alpha)					/cds=(26,790) /gb=J00194 /gi=188231 /ug=Hs.76807 /len=1199	
LRP1 (low density lipoprotein-related protein 1 (alpha-2-macroglobulin receptor))	X13916	Hs.89137	NM_002332	12q13-q14	Cluster Incl. X13916:Human mRNA for LDL-receptor related protein /cds=(466,14100) /gb=X13916 /gi=34338 /ug=Hs.89137 /len=14885	38775_at
					L47276 /FEATURE=UTR#1 /DEFINITION=HUMTOPATR Homo sapiens (cell line HL-60) alpha topoisomerase truncated-form mRNA, 3 UTR	904_s_at
CCNB1 (cyclin B1)	M25753	Hs.23960	NM_031966	5q12	Cluster Incl. M25753:Human cyclin B mRNA, 3 end /cds=UNKNOWN /gb=M25753 /gi=181243 /ug=Hs.23960 /len=1452	34736_at
FGFR1 (fibroblast growth factor receptor 1 (fms-related tyrosine kinase 2, Pfeiffer syndrome))	X66945	Hs.748	NM_000604	8p11.2-p11.1	X66945 /FEATURE=cds /DEFINITION=HSNSAMTK H.sapiens N-sam mRNA for fibroblast growth factor	424_s_at

					receptor	
TAF2D (TATA box binding protein (TBP)-associated factor, RNA polymerase II, D, 100kD)	X95525	Hs.96103	NM_006951	10q24-q25.2	Cluster Incl. X95525:H.sapiens mRNA for TAFII100 protein /cds=(23,2422) /gb=X95525 /gi=1491717 /ug=Hs.96103 /len=3089	41050_at
UBE2M (ubiquitin-conjugating enzyme E2M (homologous to yeast UBC12))	AA587372	Hs.200478	NM_003969	19q13.2	Cluster Incl. AA587372:nn82f03.s1 Homo sapiens cDNA, 3 end /clone=IMAGE-1090397 /clone_end=3 /gb=AA587372 /gi=2398186 /ug=Hs.200478 /len=636	33782_r_at
PRKCB1 (protein kinase C, beta 1)	X06318	Hs.77202	NM_002738	16p11.2	X06318 /FEATURE=cds /DEFINITION=HSPKCB1A Human mRNA for protein kinase C (PKC) type beta I	1336_s_at
AIM1 (absent in melanoma 1)	AI800499	Hs.161002		6q21	Cluster Incl. AI800499:tc11f11.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-2063565 /clone_end=3 /gb=AI800499 /gi=5365971 /ug=Hs.161002 /len=403	32112_s_at

PLSCR1 (phospholipid scramblase 1)	AB006746	Hs.198282	NM_021105	3q23	Cluster Incl. AB006746:Homo sapiens hMmTRA1b mRNA, complete cds /cds=(256,1212) /gb=AB006746 /gi=3510296 /ug=Hs.198282 /len=2077	32775_r_at
					Cluster Incl. AL080216:Homo sapiens mRNA; cDNA DKFZp586K1123 (from clone DKFZp586K1123) /cds=UNKNOWN /gb=AL080216 /gi=5262707 /ug=Hs.26837 /len=2204	35187_at
ASH2L (ash2 (absent, small, or homeotic, Drosophila, homolog)-like)	AB022785	Hs.6856	NM_004674	8p11.2	Cluster Incl. AB022785:Homo sapiens ASH2L gene, complete cds, similar to Drosophila ash2 gene /cds=(12,1898) /gb=AB022785 /gi=4210446 /ug=Hs.6856 /len=2369	35804_at
DXF68S1E(DNA segment, numerous copies, expressed probes (GS1 gene))	M86934			X	Cluster Incl. M86934:Human GS1 (protein of unknown function) mRNA, complete cds /cds=(35,679) /gb=M86934 /gi=183652 /ug=Hs.78991 /len=2058	37709_at

					Cluster Incl. AF067420:Homo sapiens SNC73 protein (SNC73) mRNA, complete cds /cds=(395,1549) /gb=AF067420 /gi=3201899 /ug=Hs.32225 /len=1594	33499_s_at
CD5 (CD5 antigen (p56-62))	X04391	Hs.58685	NM_014207	11q13	Cluster Incl. X04391:Human mRNA for lymphocyte glycoprotein T1/Leu-1 /cds=(72,1559) /gb=X04391 /gi=37186 /ug=Hs.234745 /len=2320	32953_at
					Cluster Incl. S71043:Ig alpha 2=immunoglobulin A heavy chain allotype 2 {constant region, germ line} [human, peripheral blood neutrophils, Genomic, 1799 nt] /cds=(0,1022) /gb=S71043 /gi=546798 /ug=Hs.32225 /len=1047	33500_i_at
PCF11(PCF11p homolog)	AB020631	Hs.123654	NM_015885	11	Cluster Incl. AB020631:Homo sapiens mRNA for KIAA0824 protein, partial cds /cds=(0,4936) /gb=AB020631 /gi=4240136 /ug=Hs.123654 /len=5834	41665_at

CRA(cisplatin resistance associated)	U78556	Hs.166066	NM_006697	1	U78556 /FEATURE= /DEFINITION=HSU78556 Human cisplatin resistance associated alpha protein (hCRA alpha) mRNA, complete cds	1230_g_at
PCK2 (phosphoenolpyruvate carboxykinase 2 (mitochondrial))	X92720	Hs.75812	NM_004563	14q11.2-14q21.3	Cluster Incl. X92720:H.sapiens mRNA for phosphoenolpyruvate carboxykinase /cds=(66,1988) /gb=X92720 /gi=1403049 /ug=Hs.75812 /len=2147	37188_at
					Cluster Incl. S71043:Ig alpha 2=immunoglobulin A heavy chain allotype 2 {constant region, germ line} [human, peripheral blood neutrophils, Genomic, 1799 nt] /cds=(0,1022) /gb=S71043 /gi=546798 /ug=Hs.32225 /len=1047	33501_r_at
C18orf1 (chromosome 18 open reading frame 1)	AF009425	Hs.153498	NM_004338	18p11.2	Cluster Incl. AF009425:Homo sapiens clone 22 mRNA, alternative splicing variant alpha-2, complete cds /cds=(469,1335) /gb=AF009425	40045_g_at

					/gi=2271470 /ug=Hs.153498 /len=8440	
CAPN2 (calpain 2, (m/II) large subunit)	M23254	Hs.76288	NM_001748	1q41-q42	Homo sapiens /REF=M23254 /DEF=Cluster Incl. :Human Ca2-activated neutral protease large subunit (CANP) mRNA, complete cds /cds=(130,2232) /gb= /gi=511636 /ug=Hs.76288 /len=3213 /LEN=3435	37001_at
IGL@ (immunoglobulin lambda locus)	X57809	Hs.181125		22q11.1-q11.2	Cluster Incl. X57809:Human rearranged immunoglobulin lambda light chain mRNA /cds=(114,815) /gb=X57809 /gi=33714 /ug=Hs.181125 /len=915	33273_f_at
GALT (galactose-1-phosphate uridylyltransferase)	M60091	Hs.75641	NM_000155	9p13	Cluster Incl. M60091:Homo sapiens galactose-1-phosphate uridyl transferase (GALT) mRNA, complete cds /cds=(28,1167) /gb=M60091 /gi=182950 /ug=Hs.75641 /len=1295	36664_at

APLP2 (amyloid beta (A4) precursor-like protein 2)	S60099	Hs.279518	NM_001642	11q24	Cluster Incl. S60099:APPH=amyloid precursor protein homolog [human, placenta, mRNA, 3727 nt] /cds=(72,2363) /gb=S60099 /gi=300168 /ug=Hs.64797 /len=3727	33944_at
CAST (calpastatin)	D16217	Hs.279607	NM_001750	5q14-q22	Cluster Incl. D16217:Human mRNA for calpastatin, complete cds /cds=(162,2288) /gb=D16217 /gi=303598 /ug=Hs.226067 /len=2493	41257_at
CSPG2 (chondroitin sulfate proteoglycan 2 (versican))	D32039	Hs.81800	NM_004385	5q14.3	Cluster Incl. D32039:Human pgH3 mRNA for proteoglycan PG-M(V3), complete cds /cds=(105,2072) /gb=D32039 /gi=1008912 /ug=Hs.234753 /len=2087	31682_s_at
USP9X (ubiquitin specific protease 9, X chromosome (Drosophila fat facets related))	X98296	Hs.77578	NM_004652	xp11.4	X98296 /FEATURE=cds /DEFINITION=HSUBIQHYD H.sapiens mRNA for ubiquitin hydrolase	969_s_at

SAMHD1 (SAM domain and HD domain, 1)	AL050267			20pter-q12	Cluster Incl. AL050267:Homo sapiens mRNA; cDNA DKFZp564A032 (from clone DKFZp564A032) /cds=(75,1955) /gb=AL050267 /gi=4886492 /ug=Hs.23889 /len=2195	34714_at
SHARP(Msx2 interacting nuclear target protein)	AL096858	Hs.184245	NM_015001	1	Cluster Incl. AL096858:Novel human gene mapping to chomosome 1 /cds=(331,10116) /gb=AL096858 /gi=5541864 /ug=Hs.184245 /len=11145	32172_at
KIAA0528(KIAA0528 gene product)	AB011100	Hs.30656		12	Cluster Incl. AB011100:Homo sapiens mRNA for KIAA0528 protein, complete cds /cds=(799,3507) /gb=AB011100 /gi=3043579 /ug=Hs.30656 /len=4682	35252_at
					Cluster Incl. N95443:zb81c12.s1 Homo sapiens cDNA, 3 end /clone=IMAGE-310006 /clone_end=3 /gb=N95443 /gi=1267753 /ug=Hs.19180 /len=611	33716_at

LOC56007(hypothetical protein 23851)	AF035313	Hs.10065		5	Cluster Incl. AF035313:Homo sapiens clone 23851 mRNA sequence /cds=UNKNOWN /gb=AF035313 /gi=2661075 /ug=Hs.10065 /len=1369	39517_at
KIAA0332(KIAA0332 protein)	AB002330			3	Cluster Incl. AB002330:Human mRNA for KIAA0332 gene, partial cds /cds=(0,3087) /gb=AB002330 /gi=2224604 /ug=Hs.7976 /len=6823	38030_at
CSPG2 (chondroitin sulfate proteoglycan 2 (versican))	X15998	Hs.81800	NM_004385	5q14.3	Cluster Incl. X15998:H.sapiens mRNA for the chondroitin sulphate proteoglycan versican, V1 splice-variant; precursor peptide /cds=(266,7495) /gb=X15998 /gi=37662 /ug=Hs.81800 /len=8224	38112_g_at
GAS7 (growth arrest-specific 7)	AB007854	Hs.226133	NM_003644	17p	Cluster Incl. AB007854:Homo sapiens KIAA0394 mRNA, complete cds /cds=(121,1359) /gb=AB007854 /gi=2662068 /ug=Hs.226133 /len=7979	33387_at

ICSBP1 (interferon consensus sequence binding protein 1)	M91196	Hs.14453	NM_002163	16q24	Cluster Incl. M91196:Homo sapiens DNA-binding protein mRNA, complete cds /cds=(47,1327) /gb=M91196 /gi=2275152 /ug=Hs.2286 /len=1538	32941_at
IFI30 (interferon, gamma-inducible protein 30)	J03909	Hs.14623	NM_006332	19p13.1	J03909 /FEATURE= /DEFINITION=HUMIIP Human gamma-interferon-inducible protein (IP-30) mRNA, complete cds	925_at
MGC4175(hypothetical protein MGC4175)	AI656421	Hs.322404	NM_024315	7	Cluster Incl. AI656421:tt50h10.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-2244259 /clone_end=3 /gb=AI656421 /gi=4740400 /ug=Hs.5671 /len=566	41809_at
KIAA0810(KIAA0810 protein)	AB018353	Hs.7531	NM_025154		Cluster Incl. AB018353:Homo sapiens mRNA for KIAA0810 protein, partial cds /cds=(0,2475) /gb=AB018353 /gi=3882340 /ug=Hs.7531 /len=4047	36588_at
					Cluster Incl. AI700633:we38g03.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-	34840_at

					2343412 /clone_end=3 /gb=AI700633 /gi=4988533 /ug=Hs.4815 /len=565	
AHNAK (AHNAK nucleoprotein (desmoyokin))	M80899	Hs.301417		11q12-q13	Cluster Incl. M80899:Human novel protein AHNAK mRNA, partial sequence /cds=(0,3835) /gb=M80899 /gi=178282 /ug=Hs.76549 /len=4051	37027_at
WEE1 (wee1+ (S. pombe) homolog)	W28575	Hs.75188	NM_003390	11p15.3-p15.1	Cluster Incl. W28575:51f12 Homo sapiens cDNA /gb=W28575 /gi=1308730 /ug=Hs.8151 /len=906	38102_at
MAD1L1 (MAD1 (mitotic arrest deficient, yeast, homolog)-like 1)	U33822	Hs.7345	NM_003550	7p22	U33822 /FEATURE= /DEFINITION=HSU33822 Human tax1-binding protein TXBP181 mRNA, complete cds	499_at
ECGF1 (endothelial cell growth factor 1 (platelet-derived))	M63193	Hs.73946	NM_001953	22q13.33	Cluster Incl. M63193:Human platelet-derived endothelial cell growth factor mRNA, complete cds /cds=(123,1571) /gb=M63193 /gi=189700 /ug=Hs.73946	36879_at

					/len=1587	
GBF1 (golgi-specific brefeldin A resistance factor 1)	D87435	Hs.155499	NM_004193	10q24	Cluster Incl. D87435:Human mRNA for KIAA0248 gene, partial cds /cds=(0,5077) /gb=D87435 /gi=1665764 /ug=Hs.155499 /len=5634	40123_at
RSN (restin (Reed-Steinberg cell-expressed intermediate filament-associated protein))	X64838	Hs.31638	NM_002956	12q24.3	Cluster Incl. X64838:H.sapiens mRNA for restin /cds=(132,4415) /gb=X64838 /gi=35998 /ug=Hs.31638 /len=5857	34350_at
HLA-DPB1 (major histocompatibility complex, class II, DP beta 1)	M83664	Hs.814	NM_002121	6p21.3	Cluster Incl. M83664:Human MHC class II lymphocyte antigen (HLA-DP) beta chain mRNA, complete cds /cds=(59,835) /gb=M83664 /gi=188478 /ug=Hs.814 /len=1501	38095_i_at
CAPZA2 (capping protein (actin filament) *muscle Z-line, alpha 2*)	U03851	Hs.75546	NM_006136	7q31.2-q31.3	Cluster Incl. U03851:*Human capping* protein alpha mRNA, partial cds /cds=(16,870) /gb=U03851 /gi=433307	36641_at

					/ug=Hs.75546 /len=2263	
ATP6A1 (ATPase, H+ transporting, lysosomal (vacuolar proton pump), alpha polypeptide, 70kD,	AA056747	Hs.281866	NM_001690	3p13-q13.2	Cluster Incl. AA056747:zk81f02.s1 Homo sapiens cDNA, 3 end /clone=IMAGE-489243 /clone_end=3 /gb=AA056747 /gi=1549122 /ug=Hs.5119 /len=559	34889_at
RBBP4 (retinoblastoma-binding protein 4)	X74262	Hs.16003	NM_005610	5p15.2	Cluster Incl. X74262:H.sapiens RbAp48 mRNA encoding retinoblastoma binding protein /cds=(84,1361) /gb=X74262 /gi=397375 /ug=Hs.16003 /len=2314	40418_at
KIAA0852(KIAA0852 protein)	AB020659	Hs.35276	NM_014941	22	Cluster Incl. AB020659:Homo sapiens mRNA for KIAA0852 protein, complete cds /cds=(1364,4276) /gb=AB020659 /gi=4240192 /ug=Hs.35276 /len=4467	35683_at
HLA-DQB1 (major histocompatibility complex, class II, DQ beta 1)	M60028	Hs.73931	NM_002123	6p21.3	Cluster Incl. M60028:Human MHC class II HLA-DQ-beta (DQB1,DQw9), complete cds /cds=(57,842) /gb=M60028 /gi=188114 /ug=Hs.73931 /len=1192	36878_f_at

KIAA0826(KIAA0826 protein)	AB020633	Hs.169600		4	Cluster Incl. AB020633:Homo sapiens mRNA for KIAA0826 protein, partial cds /cds=(0,3710) /gb=AB020633 /gi=4240140 /ug=Hs.169600 /len=5770	40492_at
AP3B1 (adaptor-related protein complex 3, beta 1 subunit)	U81504	Hs.155172	NM_003664	5p14.3-q14.3	Cluster Incl. U81504:Homo sapiens beta-3A-adaptin subunit of the AP-3 complex mRNA, complete cds /cds=(92,3376) /gb=U81504 /gi=2199511 /ug=Hs.155172 /len=3950	32039_at
					Cluster Incl. AL109698:Homo sapiens mRNA full length insert cDNA clone EUROIMAGE 26539 /cds=UNKNOWN /gb=AL109698 /gi=5689808 /ug=Hs.8065 /len=2035	37590_g_at
GJA8 (gap junction protein, alpha 8, 50kD (connexin 50))	U34802	Hs.157433	NM_005267	1q21.1	Cluster Incl. U34802:Human intrinsic membrane protein MP70 (Cx50) gene, complete cds /cds=(0,1298) /gb=U34802 /gi=1002998 /ug=Hs.157433 /len=1299	31778_at

JAK1 (Janus kinase 1 (a protein tyrosine kinase))	M64174	Hs.50651	NM_002227	1p32.3-p31.3	Cluster Incl. M64174:Human protein-tyrosine kinase (JAK1) mRNA, complete cds /cds=(75,3503) /gb=M64174 /gi=190734 /ug=Hs.50651 /len=3541	41594_at
CDC2 (cell division cycle 2, G1 to S and G2 to M)	D88357	Hs.184572	NM_001786	10q21.1	Cluster Incl. D88357:Homo sapiens mRNA for CDC2 delta T, complete cds /cds=(27,749) /gb=D88357 /gi=3126638 /ug=Hs.184572 /len=780	33324_s_at
NNT (nicotinamide nucleotide transhydrogenase)	U40490	Hs.18136	NM_012343	5p13.1-5cen	Cluster Incl. U40490:Human nicotinamide nucleotide transhydrogenase mRNA, nuclear gene encoding mitochondrial protein, complete cds /cds=(143,3403) /gb=U40490 /gi=1110519 /ug=Hs.18136 /len=4232	41722_at
ITPR3 (inositol 1,4,5-triphosphate receptor, type 3)	U01062	Hs.77515	NM_002224	6p21	Cluster Incl. U01062:Human type 3 inositol 1,4,5-trisphosphate receptor (ITPR3) mRNA, complete cds /cds=(36,8051) /gb=U01062 /gi=453367 /ug=Hs.77515	37343_at

					/len=8833	
KIAA0662(KIAA0662 gene product)	AB014562	Hs.93868		9	Cluster Incl. AB014562:Homo sapiens mRNA for KIAA0662 protein, partial cds /cds=(0,2034) /gb=AB014562 /gi=3327137 /ug=Hs.93868 /len=3944	39117_at
PAFAH1B1 (platelet-activating factor acetylhydrolase, isoform Ib, alpha subunit (45kD))	L13385	Hs.77318	NM_000430	17p13.3	Cluster Incl. L13385:Homo sapiens(clone 71) Miller-Dieker lissencephaly protein (LIS1) mRNA, complete cds /cds=(217,1449) /gb=L13385 /gi=349823 /ug=Hs.77318 /len=5243	32569_at
PSMD12 (proteasome (prosome, macropain) 26S subunit, non-ATPase, 12)	AB003103	Hs.4295	NM_002816	17	AB003103 /FEATURE= /DEFINITION=AB003103 Homo sapiens mRNA for 26S proteasome subunit p55, complete cds	1192_at
LY64 (lymphocyte antigen 64 (mouse) homolog, radioprotective, 105kD)	D83597	Hs.87205	NM_005582	5q12	Cluster Incl. D83597:Homo sapiens mRNA for RP105, complete cds /cds=(142,2127) /gb=D83597 /gi=1843410 /ug=Hs.87205	40715_at

/len=2697

Table 6:

UCL/HGNC/HUGO Human Gene Nomenclature Database Symbol	GenBank Accession No.	UniGene Cluster	RefSeq	Chromosomal Location	Description Unigene Build #95	
CCR2 (chemokine (C-C motif) receptor 2)	U95626	Hs.395	NM_000647	3p21	Cluster Incl. U95626:Homo sapiens ccr2b (ccr2), ccr2a (ccr2), ccr5 (ccr5) and ccr6 (ccr6) genes, complete cds, and lactoferrin (lactoferrin) gene, partial cds /cds=(2,1429) /gb=U95626 /gi=2104517 /ug=Hs.105938 /len=1607	37149_s_at
CCR2 (chemokine (C-C motif) receptor 2)	NM_000647	Hs.395	NM_000647	3p21	Cluster Incl. NM_000647 NM_000648:wi54d04.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-2394055 /clone_end=3 /gb=AI762213 /gi=5177880 /ug=Hs.204238 /len=677	32821_at
AZU1 (azurocidin 1 (cationic antimicrobial protein 37))	M96326	Hs.72885	NM_001700	19p13.3	Cluster Incl. M96326:Human azurocidin gene, complete cds /cds=(16,771)	33963_at

protein 37))					/gb=M96326 /gi=179301 /ug=Hs.72885 /len=913	
CAMP (cathelicidin antimicrobial peptide)	Z38026	Hs.51120	NM_004345	3p21.3	Cluster Incl. Z38026:H.sapiens mRNA for FALL-39 peptide antibiotic /cds=(11,523) /gb=Z38026 /gi=558378 /ug=Hs.51120 /len=615	36710_at
	D872				Cluster Incl. D872+B792:Homo sapiens mRNA for rhodanese, complete cds /cds=(48,941) /gb=D87292 /gi=1877030 /ug=Hs.74097 /len=1137	36123_at
ZWINT (ZW10 interactor)	AF067656	Hs.42650	NM_007057	10q21-q22	Cluster Incl. AF067656:Homo sapiens ZW10 interactor Zwint mRNA, complete cds /cds=(24,857) /gb=AF067656 /gi=3901271 /ug=Hs.42650 /len=1639	35995_at
CAT (catalase)	AL035079	Hs.76359	NM_001752	11p13	Cluster Incl. AL035079:dJ53C18.1 (Catalase) /cds=(74,1657) /gb=AL035079 /gi=4775614 /ug=Hs.76359 /len=2287	37009_at

NUCB2 (nucleobindin 2)	X76732	Hs.3164	NM_005013	11p15.1-p14	Cluster Incl. X76732:H.sapiens mRNA for NEFA protein /cds=(219,1481) /gb=X76732 /gi=2706486 /ug=Hs.3164 /len=1586	35643_at
RAD54L (RAD54 (S.cerevisiae)-like)	X97795	Hs.66718	NM_003579	1p32	X97795 /FEATURE=cds /DEFINITION=HSRAD54 H.sapiens mRNA homologous to S. cerevisiae RAD54	966_at
SLC29A1 (solute carrier family 29 (nucleoside transporters), member 1)	U81375	Hs.25450	NM_004955	6p21.1-p21.2	Cluster Incl. U81375:Human placental equilibrative nucleoside transporter 1 (hENT1) mRNA, complete cds /cds=(178,1548) /gb=U81375 /gi=1845344 /ug=Hs.25450 /len=2162	33901_at
[KIAA0101(KIAA0101 gene product)	D14657	Hs.81892	NM_014736	15	Cluster Incl. D14657:Human mRNA for KIAA0101 gene, complete cds /cds=(61,396) /gb=D14657 /gi=285938 /ug=Hs.81892 /len=836	38116_at

MPO (myeloperoxidase)	M19507	Hs.1817	NM_000250	17q23.1	Cluster Incl. M19507:Human myeloperoxidase mRNA, complete cds /cds=UNKNOWN /gb=M19507 /gi=188657 /ug=Hs.1817 /len=3215	33284_at
PCNA (proliferating cell nuclear antigen)	M15796	Hs.78996	NM_002592	20pter-p12	M15796 /FEATURE= /DEFINITION=HUMCYL Human cyclin protein gene, complete cds	1884_s_at
PRDX3 (peroxiredoxin 3)	D49396	Hs.75454	NM_006793	10q25-q26	Cluster Incl. D49396:Human mRNA for Apo1_Human (MER5(Aop1-Mouse)-like protein), complete cds /cds=(6,776) /gb=D49396 /gi=682747 /ug=Hs.75454 /len=1531	36631_at
IGLL1 (immunoglobulin lambda-like polypeptide 1)	M27749	Hs.288168	NM_020070	22q11.23	Cluster Incl. M27749:Human immunoglobulin-related 14.1 protein mRNA, complete cds /cds=(118,759) /gb=M27749 /gi=186145 /ug=Hs.170116 /len=847	38514_at

SNL (singed (Drosophila)-like (sea urchin fascin homolog like))	U03057	Hs.118400	NM_003088	7p22	Cluster Incl. U03057:Human actin bundling protein (HSN) mRNA, complete cds /cds=(111,1592) /gb=U03057 /gi=458027 /ug=Hs.118400 /len=2767	39070_at
TUBB2(tubulin, beta, 2)	X02344	Hs.251653	NM_006088		Cluster Incl. X02344:Homo sapiens beta 2 gene /cds=(0,1337) /gb=X02344 /gi=37493 /ug=Hs.184582 /len=1338	33679_f_at
SPINK2 (serine protease inhibitor, Kazal type, 2 (acrosin-trypsin inhibitor)	X57655	Hs.98243	NM_021114	4	Cluster Incl. X57655:H.sapiens RNA for acrosin-trypsin inhibitor (HUSI-II) /cds=(68,322) /gb=X57655 /gi=32549 /ug=Hs.98243 /len=594	41071_at
TARS (threonyl-tRNA synthetase	M63180	Hs.84131	NM_003191	5p13-cen	Cluster Incl. M63180:Human threonyl-tRNA synthetase mRNA, complete cds /cds=(138,2276) /gb=M63180 /gi=339679 /ug=Hs.84131 /len=2644	38473_at
					Rad2	1515_at

ELA2 (elastase 2, neutrophil)	M34379	Hs.99863	NM_001972	19p13.3	Cluster Incl. M34379:Human elastase/medullasin mRNA, complete cds /cds=(38,841) /gb=M34379 /gi=187116 /ug=Hs.99863 /len=920	37096_at
SCGF (stem cell growth factor; lymphocyte secreted C-type lectin)	AF020044	Hs.105927	NM_002975	19q13.3	Cluster Incl. AF020044:Homo sapiens lymphocyte secreted C-type lectin precursor, mRNA, complete cds /cds=(179,1150) /gb=AF020044 /gi=2828595 /ug=Hs.105927 /len=1391	37147_at
HNRPAB (heterogeneous nuclear ribonucleoprotein A/B)	M65028	Hs.81361	NM_004499 NM_031266	5q35	Cluster Incl. M65028:Human hnRNP type A/B protein mRNA, complete cds /cds=(142,996) /gb=M65028 /gi=337450 /ug=Hs.81361 /len=1537	38094_at
RAB32 (RAB32, member RAS oncogene family)	U59878	Hs.32217	NM_006834	6	Cluster Incl. U59878:Human low-Mr GTP-binding protein (RAB32) mRNA, partial cds /cds=(0,632) /gb=U59878 /gi=1388198 /ug=Hs.32217 /len=980	41523_at

CTSG (cathepsin G)	M16117	Hs.100764	NM_001911	14q11.2	Cluster Incl. M16117:Human cathepsin G mRNA, complete cds /cds=(8,775) /gb=M16117 /gi=181181 /ug=Hs.100764 /len=857	37105_at
H2AFY (H2A histone family, member Y)	AF054174	Hs.75258	NM_004893	5q31.3-q32	Cluster Incl. AF054174:Homo sapiens histone macroH2A1.2 mRNA, complete cds /cds=(173,1288) /gb=AF054174 /gi=3341991 /ug=Hs.75258 /len=1881	36576_at
GAPD (glyceraldehyde-3-phosphate dehydrogenase)	M33197	Hs.169476	NM_002046	12p13	Homo sapiens /REF=M33197 /DEF=Human glyceraldehyde-3-phosphate dehydrogenase (GAPDH) mRNA, complete cds /LEN=1268 (_5, _M, _3 represent transcript regions 5 prime, Middle, and 3 prime respectively)	AFFX-HUMG/
	L47276				L47276 /FEATURE=UTR#1 /DEFINITION=HUMTOPATR Homo sapiens (cell line HL-60) alpha topoisomerase truncated-form mRNA, 3	904_s_at

					UTR	
KIAA0222(KIAA0222 gene product)	AL044599	Hs.48450	NM_014643	18	Cluster Incl. AL044599:DKFZp434N192_s1 Homo sapiens cDNA, 3 end /clone=DKFZp434N192 /clone_end=3 /gb=AL044599 /gi=5432814 /ug=Hs.48450 /len=1067	34843_at
H2AFX (H2A histone family, member X)	X14850	Hs.147097	NM_002105	11q23.2-q23.3	Cluster Incl. X14850:Human H2A.X mRNA encoding histone H2A.X /cds=(73,504) /gb=X14850 /gi=31972 /ug=Hs.147097 /len=1585	40195_at
LOC94392(hypothetical gene supported by AB007931; AF055010; AK001233; AK022322;	AB007931			1	Cluster Incl. AB007931:Homo sapiens mRNA for KIAA0462 protein, partial cds /cds=(0,6831) /gb=AB007931 /gi=3413885 /ug=Hs.239686 /len=7150	33860_at
CCNB1 (cyclin B1)	M25753	Hs.23960	NM_031966	5q12	Cluster Incl. M25753:Human cyclin B mRNA, 3 end /cds=UNKNOWN	34736_at

					/gb=M25753 /gi=181243 /ug=Hs.23960 /len=1452	
LOC95295(hypothetical gene supported by V00599; BC001938; BC007605; BC008791	V00599			6	V00599 /FEATURE=mRNA /DEFINITION=HSTUB2 Human mRNA fragment encoding beta-tubulin. (from clone D-beta-1)	151_s_at
MLC1(KIAA0027 protein	D25217	Hs.74518	NM_015166	22	Cluster Incl. D25217:Human mRNA for KIAA0027 gene, partial cds /cds=(0,1317) /gb=D25217 /gi=434776 /ug=Hs.74518 /len=3435	36897_at
PRTN3 (proteinase 3 (serine proteinase, neutrophil, Wegener granulomatosis autoantigen	X55668	Hs.928	NM_002777	19p13.3	Cluster Incl. X55668:Human mRNA for proteinase 3 /cds=(0,764) /gb=X55668 /gi=35687 /ug=Hs.928 /len=965	37066_at
ALDH2 (aldehyde dehydrogenase 2 family (mitochondrial)	X05409	Hs.195432	NM_000690	12q24.2	Cluster Incl. X05409:Human RNA for mitochondrial aldehyde dehydrogenase I ALDH I (EC 1.2.1.3) /cds=(36,1586) /gb=X05409 /gi=28605 /ug=Hs.195432	32747_at

					/len=1989	
KNSL6 (kinesin-like 6 (mitotic centromere-associated kinesin)	U63743	Hs.69360	NM_006845	1	Cluster Incl. U63743:Homo sapiens mitotic centromere-associated kinesin mRNA, complete cds /cds=(54,2231) /gb=U63743 /gi=1695881 /ug=Hs.69360 /len=2740	36837_at
TST (thiosulfate sulfurtransferase (rhodanese))	X59434	Hs.248267	NM_003312	22q13.1	Cluster Incl. X59434:Human rohu mRNA for rhodanese /cds=(34,924) /gb=X59434 /gi=432375 /ug=Hs.74097 /len=1232	36124_at
PAI-RBP1(PAI-1 mRNA-binding protein	AL080119	Hs.165998	NM_015640	1	Cluster Incl. AL080119:Homo sapiens mRNA; cDNA DKFZp564M2423 (from clone DKFZp564M2423) /cds=(85,1248) /gb=AL080119 /gi=5262550 /ug=Hs.165998 /len=2183	40440_at
RAB13 (RAB13, member RAS oncogene family	X75593	Hs.151536	NM_002870	12q13	Cluster Incl. X75593:H.sapiens mRNA for rab 13 /cds=(139,750) /gb=X75593 /gi=452319 /ug=Hs.151536 /len=1238	40210_at

NCOA4 (nuclear receptor coactivator 4)	X77548	Hs.99908	NM_005437	10q11.2	Cluster Incl. X77548:H. sapiens cDNA for RFG /cds=(76,1920) /gb=X77548 /gi=469145 /ug=Hs.99908 /len=3418	39174_at
ANXA1 (annexin A1)	X05908	Hs.78225	NM_000700	9q12-q21.2	Cluster Incl. X05908:Human mRNA for lipocortin /cds=(74,1114) /gb=X05908 /gi=34387 /ug=Hs.78225 /len=1399	37403_at
FEN1 (flap structure-specific endonuclease 1)	AC004770	Hs.4756	NM_004111	11q12	Cluster Incl. AC004770:Homo sapiens chromosome 11, BAC CIT-HSP-311e8 (BC269730) containing the hFEN1 gene /cds=(2644,3786) /gb=AC004770 /gi=3212836 /ug=Hs.4756 /len=4522	41583_at
MYCBP (c-myc binding protein)	D50692	Hs.78221	NM_012333	1p33-p32.2	D50692 /FEATURE= /DEFINITION=HUMAMY1 Homo sapiens mRNA for c-myc binding protein, complete cds	1904_at
					Rad2	1516_g_at

IGFBP7 (insulin-like growth factor binding protein 7)	L19182	Hs.119206	NM_001553	4q12	L19182 /FEATURE= /DEFINITION=HUMMAC25X Human MAC25 mRNA, complete cds	2062_at
ICA1 (islet cell autoantigen 1 (69kD))	U38260	Hs.167927	NM_004968	7p22	Cluster Incl. U38260:Human islet cell autoantigen ICAp69 mRNA, complete cds /cds=(169,942) /gb=U38260 /gi=1675205 /ug=Hs.167927 /len=1415	32634_s_at
KIAA1055(KIAA1055 protein	AB028978	Hs.126084		15	Cluster Incl. AB028978:Homo sapiens mRNA for KIAA1055 protein, partial cds /cds=(0,2607) /gb=AB028978 /gi=5689446 /ug=Hs.126084 /len=5876	39400_at
CSF1 (colony stimulating factor 1 (macrophage))	M37435	Hs.173894	NM_000757	1p21-p13	M37435 /FEATURE= /DEFINITION=HUMCSDF1 Human macrophage-specific colony-stimulating factor (CSF-1) mRNA, complete cds	882_at
	W28186				Cluster Incl. W28186:43c2 Homo sapiens cDNA /gb=W28186 /gi=1308134	41188_at

					/ug=Hs.180320 /len=941	
SYNGR1 (synaptogyrin 1)	AL022326	Hs.6139	NM_004711	22q13.1	Cluster Incl. AL022326:dJ333H23.2.2 (Synaptogyrin 1A (SYNGR1A)) /cds=(43,744) /gb=AL022326 /gi=3550039 /ug=Hs.6139 /len=4406	35354_at
RNASE2 (ribonuclease, RNase A family, 2 (liver, eosinophil-derived neurotoxin)	X55988	Hs.728	NM_002934	14q24-q31	Cluster Incl. X55988:Human EDN mRNA for eosinophil derived neurotoxin /cds=(71,556) /gb=X55988 /gi=31088 /ug=Hs.728 /len=735	36766_at
S100A8 (S100 calcium-binding protein A8 (calgranulin A))	AI126134	Hs.100000	NM_002964	1q21	Cluster Incl. AI126134:qd77c05.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-1735496 /clone_end=3 /gb=AI126134 /gi=3594648 /ug=Hs.100000 /len=446	41096_at
IQGAP2 (IQ motif containing GTPase activating protein 2)	U51903	Hs.78993	NM_006633	5q	Cluster Incl. U51903:Human RasGAP-related protein (IQGAP2) mRNA, complete cds /cds=(222,4949) /gb=U51903 /gi=1262925 /ug=Hs.78993 /len=5767	37276_at

H1F0 (H1 histone family, member 0)	Z97630	Hs.226117	NM_005318	22q13.1	Cluster Incl. Z97630:Human DNA sequence from clone 466N1 on chromosome 22q12-13 Contains H1F0(H1 histone family, member 0) gene, 2-amino-3-ketobutyrate -CoA ligase(nuclear gene encoding mitochondrial protein), GALR3 (galanin receptor) gene, ESTs, GSSs and CpG islands /cds=(381,965) /gb=Z97630 /gi=4582128 /ug=Hs.226117 /len=2527	33386_at
ADAM15 (a disintegrin and metalloproteinase domain 15 (metargidin))	U41767	Hs.92208	NM_003815	1q21.3	Cluster Incl. U41767:Human metargidin precursor mRNA, complete cds /cds=(7,2451) /gb=U41767 /gi=1235673 /ug=Hs.92208 /len=2725	38282_at
AKR1C3 (aldo-keto reductase family 1, member C3 (3-alpha hydroxysteroid dehydrogenase, type	D17793	Hs.78183	NM_003739	10p15-p14	Cluster Incl. D17793:Human mRNA for KIAA0119 gene, complete cds /cds=(51,1022) /gb=D17793 /gi=457407 /ug=Hs.78183 /len=1204	37399_at
DF (D component of complement (adipsin))	M84526	Hs.155597	NM_001928	19	Cluster Incl. M84526:Human adipsin/complement factor D mRNA,	40282_s_at

					complete cds /cds=(54,740) /gb=M84526 /gi=178625 /ug=Hs.155597 /len=1071	
OAZ1 (ornithine decarboxylase antizyme 1)	D78361	Hs.125078		19p13.3	D78361 /FEATURE= /DEFINITION=HUMODAZ Human mRNA for ornithine decarboxylase antizyme, ORF 1 and ORF 2	1315_at
ERG (v-ets avian erythroblastosis virus E26 oncogene related)	M21535	Hs.45514	NM_004449	21	M21535 /FEATURE= /DEFINITION=HUMERG11 Human erg protein (ets-related gene) mRNA, complete cds	914_g_at
GPSN2 (glycoprotein, synaptic 2)	AF038958	Hs.306122	NM_004868	19p13.	Cluster Incl. AF038958:Homo sapiens synaptic glycoprotein SC2 spliced variant mRNA, complete cds /cds=(76,1002) /gb=AF038958 /gi=3329385 /ug=Hs.109051 /len=1116	38966_at
DEFA1 (defensin, alpha 1, myeloid-related sequence)	AL036554	Hs.274463	NM_004084	8p23.2-p23.1	Cluster Incl. AL036554:DKFZp564J2262_r1 Homo sapiens cDNA, 5 end /clone=DKFZp564J2262 /clone_end=5	31793_at

					/clone=DKFZp564J2262 /clone_end=5 /gb=AL036554 /gi=5927801 /ug=Hs.1379 /len=517	
MGST2 (microsomal glutathione S-transferase 2)	U77604	Hs.81874	NM_002413	4q28-q31	U77604 /FEATURE= /DEFINITION=HSU77604 Homo sapiens microsomal glutathione S-transferase 2 (MGST2) mRNA, complete cds	820_at
TOP2A (topoisomerase (DNA) II alpha (170kD))	AI375913	Hs.156346	NM_001067	17q21-q22	Cluster Incl. AI375913:tc14c08.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-2063822 /clone_end=3 /gb=AI375913 /gi=4175903 /ug=Hs.156346 /len=916	40145_at
DEFA3 (defensin, alpha 3, neutrophil-specific)	L12691	Hs.294176	NM_005217	8pter-p23.3	Cluster Incl. L12691:Human neutrophil peptide-3 gene, complete cds /cds=(50,334) /gb=L12691 /gi=292364 /ug=Hs.178741 /len=452	31506_s_at
CKS2 (CDC28 protein kinase 2)	X54942	Hs.83758	NM_001827	9q22	Cluster Incl. X54942:H.sapiens ckshs2 mRNA for Cks1 protein homologue /cds=(95,334) /gb=X54942 /gi=29978	40690_at

					/ug=Hs.83758 /len=612	
CSTA (cystatin A (stefin A))	AA570193	Hs.2621	NM_005213	3q21	Cluster Incl. AA570193:nf38c11.s1 Homo sapiens cDNA /clone=IMAGE-916052 /gb=AA570193 /gi=2344173 /ug=Hs.2621 /len=450	39581_at
HDGF (hepatoma-derived growth factor (high-mobility group protein 1-like)	L24521	Hs.89525	NM_004494	xq25	Cluster Incl. L24521:Human transformation-related protein mRNA, 3 end /cds=(0,1108) /gb=L24521 /gi=403459 /ug=Hs.169225 /len=1240	36446_s_at
TFDP2 (transcription factor Dp-2 (E2F dimerization partner 2)	L40386	Hs.19131	NM_006286	3q23	L40386 /FEATURE=mRNA /DEFINITION=HUMDP2M Human DP-2 mRNA, complete cds	633_s_at
TYMS (thymidylate synthetase)	X02308	Hs.82962	NM_001071	18p11.32	Cluster Incl. X02308:Human mRNA for thymidylate synthase (EC 2.1.1.45) /cds=(105,1046) /gb=X02308 /gi=37478 /ug=Hs.82962 /len=1536	37899_at

H2AV(histone H2A.F/Z variant	AW007731	Hs.301005	NM_012412	7	Cluster Incl. AW007731:wt68d11.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-2512629 /clone_end=3 /gb=AW007731 /gi=5856509 /ug=Hs.9242 /len=659	39092_at
EPB72 (erythrocyte membrane protein band 7.2 (stomatin)	X85116	Hs.160483	NM_004099	9q34.1	Cluster Incl. X85116:H.sapiens epb72 gene exon 1 /cds=(61,927) /gb=X85116 /gi=1161561 /ug=Hs.160483 /len=3035	40419_at
GNAQ (guanine nucleotide binding protein (G protein), q polypeptide	U40038	Hs.296261	NM_002072	9q21	Cluster Incl. U40038:Human GTP-binding protein alpha q subunit (GNAQ) mRNA, complete cds /cds=(42,1121) /gb=U40038 /gi=1181670 /ug=Hs.180950 /len=1450	38581_at
HBD (hemoglobin, delta)	V00505	Hs.36977	NM_000519	11p15.5	Cluster Incl. V00505:Human gene for delta-globin /cds=(50,493) /gb=V00505 /gi=30510 /ug=Hs.36977 /len=624	33516_at
TTK (TTK protein kinase)	M86699	Hs.169840	NM_003318	6q13-q21	M86699 /FEATURE= /DEFINITION=HUMTTK Human kinase (TTK) mRNA, complete cds	572_at

KIAA0661(95 kDa retinoblastoma protein binding protein	AB014561	Hs.65238	NM_014771	16	Cluster Incl. AB014561:Homo sapiens mRNA for KIAA0661 protein, complete cds /cds=(92,3097) /gb=AB014561 /gi=3327135 /ug=Hs.65238 /len=4199	35768_at
MCM3 (minichromosome maintenance deficient (S. cerevisiae) 3	D38073	Hs.179565	NM_002388	6p12	Cluster Incl. D38073:Human mRNA for hRlf beta subunit (p102 protein), complete cds /cds=(77,2503) /gb=D38073 /gi=862331 /ug=Hs.179565 /len=3071	33252_at
KIAA0161(ubiquitin conjugating enzyme 7 interacting protein 4	D79983	Hs.78894	NM_014746	2	Cluster Incl. D79983:Human mRNA for KIAA0161 gene, complete cds /cds=(348,1226) /gb=D79983 /gi=1136383 /ug=Hs.78894 /len=5559	37695_at
CCNB1 (cyclin B1	M25753	Hs.23960	NM_031966	5q12	M25753 /FEATURE=mRNA /DEFINITION=HUMCYCB Human cyclin B mRNA, 3 end	1945_at
PK428(Ser-Thr protein kinase related to the myotonic dystrophy protein kinase	U59305	Hs.44708	NM_003607	1	Cluster Incl. U59305:Human ser-thr protein kinase PK428 mRNA, complete cds /cds=(1288,2778) /gb=U59305	39962_at

					/gi=1695872 /ug=Hs.44708 /len=2785	
MYB (v-myb avian myeloblastosis viral oncogene homolog	M15024	Hs.1334	NM_005375	6q22-q23	M15024 /FEATURE= /DEFINITION=HUMCMYBLA Human c-myb mRNA, complete cds	2042_s_at
OAT (ornithine aminotransferase (gyrate atrophy)	M12267	Hs.75485	NM_000274	10q26	Cluster Incl. M12267:Human ornithine aminotransferase mRNA, complete cds /cds=(54,1373) /gb=M12267 /gi=189328 /ug=Hs.75485 /len=2013	36636_at
LDHA (lactate dehydrogenase A)	X02152	Hs.2795	NM_005566	11p15.4	Cluster Incl. X02152:Human mRNA for lactate dehydrogenase-A (LDH-A, EC 1.1.1.27) /cds=(97,1095) /gb=X02152 /gi=34312 /ug=Hs.2795 /len=1661	41485_at
P311(P311 protein)	U30521	Hs.142827	NM_004772	8	Cluster Incl. U30521:Human P311 HUM (3.1) mRNA, complete cds /cds=(202,408) /gb=U30521 /gi=963091 /ug=Hs.142827 /len=2036	39710_at

AGPS (alkylglycerone phosphate synthase	Y09443	Hs.22580	NM_003659	2q31	Cluster Incl. Y09443:H.sapiens mRNA for alkyl-dihydroxyacetonephosphate synthase precursor /cds=(15,1991) /gb=Y09443 /gi=1922284 /ug=Hs.22580 /len=2074	39225_at
GSN (gelsolin (amyloidosis, Finnish type)	X04412	Hs.290070	NM_000177	9q33	Cluster Incl. X04412:Human mRNA for plasma gelsolin - /cds=(14,2362) /gb=X04412 /gi=35447 /ug=Hs.80562 /len=2602	32612_at
ALDH3B1 (aldehyde dehydrogenase 3 family, member B1	U10868	Hs.83155	NM_000694	11q13	Cluster Incl. U10868:Human aldehyde dehydrogenase ALDH7 mRNA, complete cds /cds=(47,1453) /gb=U10868 /gi=601779 /ug=Hs.83155 /len=2790	40685_at
TFDP1 (transcription factor Dp-1	L23959	Hs.79353	NM_007111	13q34	L23959 /FEATURE= /DEFINITION=HUMDP1A Homo sapiens E2F-related transcription factor (DP-1) mRNA, complete cds	1670_at

TUBG1 (tubulin, gamma 1)	M61764	Hs.21635	NM_001070	17q21-q22	Cluster Incl. M61764:Human gamma-tubulin mRNA, complete cds /cds=(24,1379) /gb=M61764 /gi=183702 /ug=Hs.21635 /len=1568	33346_r_at
LMO2 (LIM domain only 2 (rhombotin-like 1)	X61118	Hs.184585	NM_005574	11p13	Cluster Incl. X61118:Human TTG-2 mRNA for a cysteine rich protein with LIM motif /cds=UNKNOWN /gb=X61118 /gi=663012 /ug=Hs.184585 /len=2292	32184_at
TPM1 (tropomyosin 1 (alpha)	M19267	Hs.77899	NM_000366	15q22.1	Cluster Incl. M19267:Human tropomyosin mRNA, complete cds /cds=(286,1140) /gb=M19267 /gi=339943 /ug=Hs.77899 /len=1633	36791_g_at
VBP1 (von Hippel-Lindau binding protein 1)	U56833	Hs.198307	NM_003372	xq28	U56833 /FEATURE= /DEFINITION=HSU56833 Human VHL binding protein-1 (VBP-1) mRNA, partial cds	171_at
TK1 (thymidine kinase 1, soluble	K02581	Hs.105097	NM_003258	17q23.2-q25.3	Cluster Incl. K02581:Human thymidine kinase mRNA, complete cds /cds=(57,761)	41400_at

					/gb=K02581 /gi=339708 /ug=Hs.105097 /len=1421	
HPRT1 (hypoxanthine phosphoribosyltransferase 1 (Lesch-Nyhan syndrome)	M31642	Hs.82314	NM_000194	xq26.1	Cluster Incl. M31642:Human hypoxanthine phosphoribosyltransferase (HPRT) mRNA, complete cds /cds=(85,741) /gb=M31642 /gi=184349 /ug=Hs.82314 /len=1331	37640_at
TUBB (tubulin, beta polypeptide	J00314	Hs.336780	NM_001069	6p21.3	J00314 /FEATURE=mRNA#1 /DEFINITION=HUMTBBM40 Human beta-tubulin gene, clone m40	709_at
SERPINB1 (serine (or cysteine) proteinase inhibitor, clade B (ovalbumin), member 1)	M93056	Hs.183583	NM_030666	6p25	Cluster Incl. M93056:Human mononcyte/neutrophil elastase inhibitor mRNA sequence /cds=UNKNOWN /gb=M93056 /gi=188621 /ug=Hs.183583 /len=1298	33305_at
MAPKAPK3 (mitogen-activated protein kinase-activated protein kinase 3)	U09578	Hs.227789	NM_004635	3p21.3	U09578 /FEATURE= /DEFINITION=HSU09578 Homo sapiens MAPKAP kinase (3pK) mRNA, complete	1637_at

					cds	
KNSL2 (kinesin-like 2)	D14678	Hs.20830		6p21.3	D14678 /FEATURE= /DEFINITION=HUMMHCB Human mRNA for kinesin-related protein, partial cds	348_at
KNSL1 (kinesin-like 1)	U37426	Hs.8878	NM_004523	10q24.1	Cluster Incl. U37426:Human kinesin-like spindle protein HKSP (HKSP) mRNA, complete cds /cds=(90,3260) /gb=U37426 /gi=1171152 /ug=Hs.8878 /len=4858	40726_at
MCM6 (minichromosome maintenance deficient (mis5, S. pombe) 6	D84557	Hs.155462	NM_005915	2q21	Cluster Incl. D84557:Homo sapiens mRNA for HsMcm6, complete cds /cds=(61,2526) /gb=D84557 /gi=1944481 /ug=Hs.155462 /len=2917	40117_at
KCNAB2 (potassium voltage-gated channel, shaker-related subfamily, beta member 2	AF044253	Hs.298184	NM_003636	1p36.3	Cluster Incl. AF044253:Homo sapiens potassium channel beta 2 subunit (HKvbeta2.2) mRNA, alternatively spliced, complete cds /cds=(0,1061) /gb=AF044253 /gi=2827465	31901_at

					/ug=Hs.154417 /len=1062	
KRT10 (keratin 10 (epidermolytic hyperkeratosis; keratosis palmaris et plantaris))	X14487	Hs.99936	NM_000421	17q21-q23	Cluster Incl. X14487:Human gene for acidic (type I) cytokeratin 10 /cds=(25,1806) /gb=X14487 /gi=28316 /ug=Hs.99936 /len=2166	38610_s_at
TNFRSF7 (tumor necrosis factor receptor superfamily, member 7)	M63928	Hs.180841	NM_001242	12p13	Cluster Incl. M63928:Homo sapiens T cell activation antigen (CD27) mRNA, complete cds /cds=(100,882) /gb=M63928 /gi=180084 /ug=Hs.180841 /len=1204	38578_at
NCOA1 (nuclear receptor coactivator 1)	AJ000882	Hs.74002	NM_003743	2p23	Cluster Incl. AJ000882:Homo sapiens mRNA for steroid receptor coactivator 1e /cds=(201,4400) /gb=AJ000882 /gi=2924310 /ug=Hs.74002 /len=4709	36118_at
SNRPN (small nuclear ribonucleoprotein polypeptide N)	U41303	Hs.48375	NM_022807	15q12	Cluster Incl. U41303:Human small nuclear ribonuleoprotein particle N (SNRPN) mRNA, complete cds /cds=(465,1187) /gb=U41303 /gi=1145774 /ug=Hs.48375	34842_at

					/len=1326	
NCOA1 (nuclear receptor coactivator 1)	U59302	Hs.74002	NM_003743	2p23	U59302 /FEATURE= /DEFINITION=HSU59302 Human steroid receptor coactivator-1 F-SRC-1 mRNA, complete cds	484_at
CDW52 (CDW52 antigen (CAMPATH-1 antigen))	N90866	Hs.276770	NM_001803	1p36	Cluster Incl. N90866:zb11b10.s1 Homo sapiens cDNA, 3 end /clone=IMAGE-301723 /clone_end=3 /gb=N90866 /gi=1444193 /ug=Hs.214742 /len=577	34210_at
SSH3BP1 (spectrin SH3 domain binding protein 1)	AF001628	Hs.24752	NM_005470	10p11.2	Cluster Incl. AF001628:Homo sapiens interactor protein AbIBP4 (AbIBP4) mRNA, complete cds /cds=(48,1403) /gb=AF001628 /gi=4100618 /ug=Hs.204036 /len=2175	38924_s_at
PKD2 (polycystic kidney disease 2 (autosomal dominant))	AL050147	Hs.91146	NM_016457	19q13.2	Cluster Incl. AL050147:Homo sapiens mRNA; cDNA DKFZp586E0820 (from clone DKFZp586E0820) /cds=(0,1630)	38269_at

					/gb=AL050147 /gi=4884153 /ug=Hs.91146 /len=1837	
MAPK3 (mitogen-activated protein kinase 3)	X60188	Hs.861		16p12-p11.2	X60188 /FEATURE=mRNA /DEFINITION=HSERK1 Human ERK1 mRNA for protein serine/threonine kinase	1000_at
HLA-DMA (major histocompatibility complex, class II, DM alpha)	X62744	Hs.77522	NM_006120	6p21.3	Cluster Incl. X62744:Human RING6 mRNA for HLA class II alpha chain-like product /cds=(45,830) /gb=X62744 /gi=36062 /ug=Hs.77522 /len=1079	37344_at
	AF038199				Cluster Incl. AF038199:Homo sapiens clone 23728 mRNA sequence /cds=UNKNOWN /gb=AF038199 /gi=2795920 /ug=Hs.153106 /len=1112	38154_at
FCGR2B (Fc fragment of IgG, low affinity IIb, receptor for (CD32))	M28696	Hs.278443	NM_004001	1q23	Cluster Incl. M28696:Human low-affinity IgG Fc receptor (beta-Fc-gamma-RII) mRNA, complete cds /cds=(41,916) /gb=M28696 /gi=184843 /ug=Hs.233450	34663_at

					/len=1416	
LYN (v-yes-1 Yamaguchi sarcoma viral related oncogene homolog)	M16038	Hs.80887	NM_002350	8q13	Cluster Incl. M16038:Human lyn mRNA encoding a tyrosine kinase /cds=(297,1835) /gb=M16038 /gi=187268 /ug=Hs.80887 /len=2298	32616_at
CELSR1 (cadherin, EGF LAG seven-pass G-type receptor 1, flamingo (Drosophila) homolog)	AL031588	Hs.252387	NM_014246	22q13.3	Cluster Incl. AL031588:dJ1163J1.1 (ortholog of mouse transmembrane receptor Celsr1 (KIAA0279 LIKE EGF-like domain containing protein similar to rat MEG /cds=(0,4433) /gb=AL031588 /gi=4007108 /ug=Hs.123043 /len=6438	41660_at
LRMP (lymphoid-restricted membrane protein)	U10485	Hs.40202	NM_006152	12p12	Cluster Incl. U10485:Human lymphoid-restricted membrane protein (Jaw1) mRNA, complete cds /cds=(574,2241) /gb=U10485 /gi=505685 /ug=Hs.40202 /len=2417	35974_at

NAF1(Nef-associated factor 1)	AJ011896	Hs.109281	NM_006058	5	Cluster Incl. AJ011896:Homo sapiens mRNA for HIV-1, Nef-associated factor 1 beta (Naf1 beta) /cds=(110,2017) /gb=AJ011896 /gi=3758820 /ug=Hs.109281 /len=2710	38970_s_at
KIAA1002(KIAA1002 protein)	AB023219	Hs.20340			Cluster Incl. AB023219:Homo sapiens mRNA for KIAA1002 protein, complete cds /cds=(800,3322) /gb=AB023219 /gi=4589647 /ug=Hs.102483 /len=4331	41366_at
DKFZP434C171(DKFZP434C171 protein)	AL080169	Hs.209100	NM_015621	5	Cluster Incl. AL080169:Homo sapiens mRNA; cDNA DKFZp434C171 (from clone DKFZp434C171) /cds=(0,544) /gb=AL080169 /gi=5262637 /ug=Hs.209100 /len=2595	34183_at
SEP2(septin 6)	D50918	Hs.90998	NM_015129	X	Cluster Incl. D50918:Human mRNA for KIAA0128 gene, partial cds /cds=(0,1276) /gb=D50918 /gi=1469178 /ug=Hs.90998 /len=4612	38826_at

PFTK1 (PFTAIRE protein kinase 1)	AB020641	Hs.57856	NM_012395	7q21-q22	Cluster Incl. AB020641:Homo sapiens mRNA for KIAA0834 protein, complete cds /cds=(144,1499) /gb=AB020641 /gi=4240156 /ug=Hs.57856 /len=4957	36502_at
DCTD (dCMP deaminase)	L39874	Hs.76894	NM_001921	4	L39874 /FEATURE=expanded_cds /DEFINITION=HUMDODDA Homo sapiens deoxycytidylate deaminase gene, complete cds	631_g_at
SQV7L(nucleotide-sugar transporter similar to C. elegans sqv-7)	AJ005866	Hs.90078		9	Cluster Incl. AJ005866:Homo sapiens mRNA for putative Sqv-7-like protein, partial /cds=(0,785) /gb=AJ005866 /gi=4008516 /ug=Hs.90078 /len=1321	38005_at
SH3GLB1 (SH3-domain, GRB2-like, endophilin B1)	AB007960	Hs.136309	NM_016009	1p22	Cluster Incl. AB007960:chromosome 1 specific transcript KIAA0491 /cds=UNKNOWN /gb=AB007960 /gi=3413934 /ug=Hs.136309 /len=5717	39691_at

PPP1CC (protein phosphatase 1, catalytic subunit, gamma isoform)	X74008	Hs.79081	NM_002710	12q24.1-q24.2	Homo sapiens /REF=X74008 /DEF=Cluster Incl. :H.sapiens mRNA for protein phosphatase 1 gamma /cds=(154,1125) /gb= /gi=402777 /ug=Hs.79081 /len=2263 /LEN=2431	37725_at
CD37 (CD37 antigen)	X14046	Hs.153053	NM_001774	19p13-q13.4	Cluster Incl. X14046:Human mRNA for leukocyte antigen CD37 /cds=(63,908) /gb=X14046 /gi=29793 /ug=Hs.153053 /len=1125	31870_at
HLA-DMB (major histocompatibility complex, class II, DM beta)	U15085	Hs.1162	NM_002118	6p21.3	Cluster Incl. U15085:Human HLA-DMB mRNA, complete cds /cds=(233,1024) /gb=U15085 /gi=557701 /ug=Hs.1162 /len=1362	41609_at
LAMA5 (laminin, alpha 5)	AB011105	Hs.11669	NM_005560	20q13.2-q13.3	Cluster Incl. AB011105:Homo sapiens mRNA for KIAA0533 protein, partial cds /cds=(0,4939) /gb=AB011105 /gi=3043589 /ug=Hs.11669 /len=5117	41610_at

EIF4B (eukaryotic translation initiation factor 4B)	X55733	Hs.93379	NM_001417	12q13.11-12q14.3	Cluster Incl. X55733:H.sapiens initiation factor 4B cDNA /cds=(0,1835) /gb=X55733 /gi=288099 /ug=Hs.93379 /len=1836	39110_at
PRKCB1 (protein kinase C, beta 1)	X07109	Hs.77202	NM_002738	16p11.2	X07109 - /FEATURE=cds /DEFINITION=HSPKCB2A Human mRNA for protein kinase C (PKC) type beta II /NOTE=replacement of probe set 1216_at	160029_at
HBOA(histone acetyltransferase)	AI951946	Hs.21907	NM_007067	X	Cluster Incl. AI951946:wx39f10.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-2546059 /clone_end=3 /gb=AI951946 /gi=5744256 /ug=Hs.244 /len=523	41338_at
CDC10 (CDC10 (cell division cycle 10, S. cerevisiae, homolog))	S72008	Hs.184326	NM_001788	7p14.3-p14.1	Cluster Incl. S72008:hCDC10=CDC10 homolog [human, fetal lung, mRNA, 2314 nt] /cds=(48,1304) /gb=S72008 /gi=560622 /ug=Hs.184326 /len=2314	32175_at
KIAA0226(KIAA0226 gene product)	D86979	Hs.141296		3	Cluster Incl. D86979:Human mRNA for KIAA0226 gene, complete cds /cds=(622,2877) /gb=D86979 /gi=1504031	31802_at

					/ug=Hs.141296 /len=5891	
BLK (B lymphoid tyrosine kinase)	S76617	Hs.2243	NM_001715	8p23-p22	S76617 /FEATURE= /DEFINITION=S76617 blk=protein tyrosine kinase [human, B lymphocytes, mRNA, 2608 nt]	854_at
HLA-DQB1 (major histocompatibility complex, class II, DQ beta 1)	M81141	Hs.73931	NM_002123	6p21.3	Cluster Incl. M81141:Human MHC class II HLA-DQ-beta mRNA (DR7 DQw2), complete cds /cds=(35,820) /gb=M81141 /gi=188202 /ug=Hs.73933 /len=1171	36773_f_at
DKFZP564K0822(hypothetical protein DKFZp564K0822)	W25986	Hs.4750	NM_030796	7	Cluster Incl. W25986:17e7 Homo sapiens cDNA /gb=W25986 /gi=1306253 /ug=Hs.4750 /len=769	34830_at
PIM2 (pim-2 oncogene)	U77735	Hs.80205	NM_006875	X	U77735 /FEATURE= /DEFINITION=HSU77735 Human pim-2 protooncogene homolog pim-2h mRNA, complete cds	1633_g_at

FCER2 (Fc fragment of IgE, low affinity II, receptor for (CD23A))	M15059	Hs.1416	NM_002002	19p13.3	Cluster Incl. M15059:Human Fc-epsilon receptor (IgE receptor) mRNA, complete cds (H107 epitope) /cds=(213,1178) /gb=M15059 /gi=182447 /ug=Hs.1416 /len=1530	34960_g_at
UBE2G2 (ubiquitin-conjugating enzyme E2G 2 (homologous to yeast UBC7))	AF032456	Hs.192853	NM_003343	21q22.3	Cluster Incl. AF032456:Homo sapiens ubiquitin conjugating enzyme G2 (UBE2G2) mRNA, complete cds /cds=(55,552) /gb=AF032456 /gi=3004908 /ug=Hs.192853 /len=2890	32236_at
SH3BP5 (SH3-domain binding protein 5 (BTK-associated))	AB005047	Hs.109150	NM_004844	1q43	Cluster Incl. AB005047:Homo sapiens mRNA for SH3 binding protein, complete cds /cds=(63,1340) /gb=AB005047 /gi=3116213 /ug=Hs.109150 /len=2570	38968_at
TRIP3 (thyroid hormone receptor interactor 3)	L40410	Hs.2210	-	17	Cluster Incl. L40410:Homo sapiens thyroid receptor interactor (TRIP3) mRNA, 3 end of cds /cds=(0,458) /gb=L40410 /gi=703109 /ug=Hs.2210 /len=867	41251_at

DKFZP586F2423(hypothetical protein DKFZp586F2423)	AL080209	Hs.13659		7	Cluster Incl. AL080209:Homo sapiens mRNA; cDNA DKFZp586F2423 (from clone DKFZp586F2423) /cds=UNKNOWN /gb=AL080209 /gi=5262698 /ug=Hs.13659 /len=4241	39692_at
KIAA0911(calsyntenin 1)	AB020718	Hs.29665	NM_014944	1	Cluster Incl. AB020718:Homo sapiens mRNA for KIAA0911 protein, complete cds /cds=(793,3738) /gb=AB020718 /gi=4240310 /ug=Hs.29665 /len=5219	41498_at
UBE2N (ubiquitin-conjugating enzyme E2N (homologous to yeast UBC13))	D83004	Hs.75355	NM_003348	12	Cluster Incl. D83004:Human epidermoid carcinoma mRNA for ubiquitin-conjugating enzyme E2 similar to Drosophila bendless gene product, complete cds /cds=(63,521) /gb=D83004 /gi=1181557 /ug=Hs.75355 /len=1203	36604_at
KIAA0542(KIAA0542 gene product)	AB011114	Hs.62209		22	Cluster Incl. AB011114:Homo sapiens mRNA for KIAA0542 protein, complete cds /cds=(393,3299) /gb=AB011114	36545_s_at

					/gi=3043607 /ug=Hs.62209 /len=5280	
					Transcription Factor Oct-1a/1b, Alt. Splice 2, Oct-1b	1171_s_at
SWAP2(suppressor of white apricot homolog 2)	AF042800	Hs.43543	NM_007056	19	Cluster Incl. AF042800:Homo sapiens suppressor of white apricot homolog 2 (SWAP2) mRNA, complete cds /cds=(143,2122) /gb=AF042800 /gi=3941325 /ug=Hs.43543 /len=2233	36005_at
MSF (MLL septin-like fusion (NOTE: non-standard symbol and name))	AB023208	Hs.181002	NM_006640	17q25	Cluster Incl. AB023208:Homo sapiens mRNA for KIAA0991 protein, complete cds /cds=(732,2000) /gb=AB023208 /gi=4589625 /ug=Hs.181002 /len=3938	41220_at
IGHM (immunoglobulin heavy constant mu)	X67301	Hs.302063		14q32.33	Cluster Incl. X67301:H.sapiens mRNA for IgM heavy chain constant region (Ab63) /cds=(0,1361) /gb=X67301 /gi=38407 /ug=Hs.179543 /len=1453	41165_g_at

					Cluster Incl. AI700633:we38g03.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-2343412 /clone_end=3 /gb=AI700633 /gi=4988533 /ug=Hs.4815 /len=565	34840_at
IGHM (immunoglobulin heavy constant mu)	X67301	Hs.302063		14q32.33	Cluster Incl. X67301:H.sapiens mRNA for IgM heavy chain constant region (Ab63) /cds=(0,1361) /gb=X67301 /gi=38407 /ug=Hs.179543 /len=1453	41164_at
ITGB7 (integrin, beta 7)	M68892	Hs.1741	NM_000889	12q13.13	M68892 /FEATURE= /DEFINITION=HUMINTB7 Human integrin beta-7 subunit mRNA, complete cds	2019_s_at
CD19 (CD19 antigen)	M28170	Hs.96023	NM_001770	16p11.2	M28170 /FEATURE= /DEFINITION=HUMCSPC Human cell surface protein CD19 (CD19) gene, complete cds	1096_g_at
PRKRIR (protein-kinase, interferon-inducible double stranded RNA dependent inhibitor, repressor of (P58 repressor))	AL049970	Hs.177574	NM_004705	11q13.5	Cluster Incl. AL049970:Homo sapiens mRNA; cDNA DKFZp564B102 (from clone DKFZp564B102) /cds=(0,965)	41141_at

repressor of (P58 repressor))					/gb=AL049970 /gi=4884219 /ug=Hs.177574 /len=2724	
NIFU(nitrogen fixation cluster-like)	U47101	Hs.9908		12	Cluster Incl. U47101:Human NifU-like protein (hNifU) mRNA, partial cds /cds=(0,366) /gb=U47101 /gi=1685101 /ug=Hs.9908 /len=819	39165_at
UBE2D2 (ubiquitin-conjugating enzyme E2D 2 (homologous to yeast UBC4/5))	AI310002	Hs.108332	NM_003339	5p14.2-q23.3	Cluster Incl. AI310002:qo77c11.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-1914548 /clone_end=3 /gb=AI310002 /gi=4004873 /ug=Hs.108332 /len=656	38705_at
	AB018272				Cluster Incl. AB018272:Homo sapiens mRNA for KIAA0729 protein, partial cds /cds=(0,3591) /gb=AB018272 /gi=3882178 /ug=Hs.180948 /len=4143	41218_at
PSCD1 (pleckstrin homology, Sec7 and coiled/coil domains 1(cytohesin 1))	M85169	Hs.1050	NM_004762	17q25	Cluster Incl. M85169:Human homologue of yeast sec7 mRNA, complete cds /cds=(69,1265) /gb=M85169 /gi=338001	38666_at

					/ug=Hs.1050 /len=3301	
LYN (v-yes-1 Yamaguchi sarcoma viral related oncogene homolog)	M16038	Hs.80887	NM_002350	8q13	M16038 /FEATURE= /DEFINITION=HUMLYN Human lyn mRNA encoding a tyrosine kinase	1402_at
HLA-DPA1 (major histocompatibility complex, class II, DP alpha 1)	X00457	Hs.914		6p21.3	Cluster Incl. X00457:Human mRNA for SB classII histocompatibility antigen alpha-chain /cds=(0,702) /gb=X00457 /gi=36405 /ug=Hs.914 /len=1048	38833_at
IGHM (immunoglobulin heavy constant mu)	X58529	Hs.302063		14q32.33	Cluster Incl. X58529:Human rearranged immunoglobulin mRNA for mu heavy chain enhancer and constant region /cds=UNKNOWN /gb=X58529 /gi=33480 /ug=Hs.179543 /len=2325	41166_at
BCL11A (B-cell CLL/lymphoma 11A (zinc finger protein))	W27619	Hs.130881	NM_022893	2p24.3-p24.1	Cluster Incl. W27619:35c7 Homo sapiens cDNA /gb=W27619 /gi=1307567 /ug=Hs.25816 /len=674	41356_at

KIAA0663(KIAA0663 gene product)	AB014563	Hs.17969	NM_014827	1	Cluster Incl. AB014563:Homo sapiens mRNA for KIAA0663 protein, complete cds /cds=(213,2645) /gb=AB014563 /gi=3327139 /ug=Hs.17969 /len=4365	41170_at
TAB2(TAK1-binding protein 2)	AB018276	Hs.109727	NM_015093	6	Cluster Incl. AB018276:Homo sapiens mRNA for KIAA0733 protein, partial cds /cds=(0,1586) /gb=AB018276 /gi=3882186 /ug=Hs.109727 /len=3479	38980_at
SETBP1 (SET binding protein 1)	AB022660	Hs.151717	NM_015559	18q21.1	Cluster Incl. AB022660:Homo sapiens mRNA for SET-binding protein (SEB), complete cds /cds=(5,4633) /gb=AB022660 /gi=5478317 /ug=Hs.151717 /len=5744	34990_at
JAK1 (Janus kinase 1 (a protein tyrosine kinase))	AL039831	Hs.50651	NM_002227	1p32.3-p31.3	Cluster Incl. AL039831:DKFZp434D1112_s1 Homo sapiens cDNA, 3 end /clone=DKFZp434D1112 /clone_end=3 /gb=AL039831 /gi=5866713 /ug=Hs.50651	34877_at

					/len=579	
ADPRTL3 (ADP-ribosyltransferase (NAD+; poly (ADP-ribose) polymerase)-like 3)	AL050034	Hs.271742	NM_005485	3p22.2-p21.1	Cluster Incl. AL050034:Homo sapiens mRNA; cDNA DKFZp566G0224 (from clone DKFZp566G0224) /cds=(0,1380) /gb=AL050034 /gi=4884274 /ug=Hs.33573 /len=1762	39670_at
IGBP1 (immunoglobulin (CD79A) binding protein 1)	Y08915	Hs.3631	NM_001551	xq13.1-q13.3	Cluster Incl. Y08915:H.sapiens mRNA for alpha 4 protein /cds=(8,1027) /gb=Y08915 /gi=1877201 /ug=Hs.3631 /len=1321	34391_at
S100A1 (S100 calcium-binding protein A1)	X58079	Hs.292707	NM_006271	1q21	Cluster Incl. X58079:Human mRNA for S100 alpha protein /cds=(113,397) /gb=X58079 /gi=36175 /ug=Hs.234348 /len=594	34674_at
HLA-DRB1 (major histocompatibility complex, class II, DR beta 1)	M32578	Hs.180255	NM_002124	6p21.3	Cluster Incl. M32578:Human MHC class II HLA-DR beta-1 mRNA (DR2.3), 5end /cds=(61,881) /gb=M32578 /gi=188305	41723_s_at

					/ug=Hs.181366 /len=1216	
SP140(nuclear body protein Sp140)	U36500	Hs.309943	NM_007237	2	Cluster Incl. U36500:Human lymphoid-specific SP100 homolog (LYSP100-B) mRNA, complete cds /cds=(116,2764) /gb=U36500 /gi=1173653 /ug=Hs.85283 /len=3252	40700_at
NCOA3 (*nuclear receptor coactivator* 3)	AF012108	Hs.225977	NM_006534	20q12	Cluster Incl. AF012108:Homo sapiens Amplified in Breast Cancer (AIB1) mRNA, complete cds /cds=(200,4462) /gb=AF012108 /gi=2331249 /ug=Hs.225977 /len=6818	33381_at
TRIAD3(TRIAD3 protein)	AA650210	Hs.86228	NM_019011	7	Cluster Incl. AA650210:ns88b12.s1 Homo sapiens cDNA /clone=IMAGE-1190687 /gb=AA650210 /gi=2577538 /ug=Hs.116406 /len=528	37476_at
ZNF9 (zinc finger protein 9 (a cellular retroviral nucleic acid binding protein))	U19765	Hs.2110	NM_003418	3q13.3-q24	Cluster Incl. U19765:Human nucleic acid binding protein gene, complete cds /cds=(14,547) /gb=U19765 /gi=790570	32841_at

					/ug=Hs.2110 /len=1665	
APOC4 (apolipoprotein C-IV)	U32576	Hs.110675	NM_001646	19q13.2	Cluster Incl. U32576:Human apolipoprotein apoC-IV (APOC4) gene, complete cds /cds=(40,423) /gb=U32576 /gi=975892 /ug=Hs.110675 /len=613	34454_r_at
CBX7 (chromobox homolog 7)	AL031846			22q13.1	Cluster Incl. AL031846:dJ742C19.5 (novel Chromobox protein) /cds=(89,844) /gb=AL031846 /gi=4164368 /ug=Hs.7442 /len=3964	36894_at
	W30677				Cluster Incl. W30677:zb75h10.r1 Homo sapiens cDNA, 5 end /clone=IMAGE-309475 /clone_end=5 /gb=W30677 /gi=1311730 /ug=Hs.5019 /len=614	34871_at
IL2RB (*interleukin 2 receptor, beta*)	AL022314	Hs.75596	NM_000878	22q13.1	Cluster Incl. AL022314:dJ1170K4.1 (novel protein similar to KIAA0176 and mouse, worm and fly proteins) /cds=(185,1057) /gb=AL022314 /gi=4090209 /ug=Hs.94810	41036_at

					/len=1854	
RERE (arginine-glutamic acid dipeptide (RE) repeats)	AB007927	Hs.194369	NM_012102	1p36.1-p36.2	Cluster Incl. AB007927:Homo sapiens mRNA for KIAA0458 protein, complete cds /cds=(155,3961) /gb=AB007927 /gi=3413877 /ug=Hs.194369 /len=6642	32253_at
CD79A (CD79A antigen (immunoglobulin-associated alpha))	U05259	Hs.79630	NM_001783	19q13.2	Cluster Incl. U05259:Human MB-1 gene, complete cds /cds=(36,716) /gb=U05259 /gi=452561 /ug=Hs.79630 /len=1107	38017_at
PLCG1 (phospholipase C, gamma 1 (formerly subtype 148))	AL022394	Hs.268177	NM_002660	20q12-q13.1	Cluster Incl. AL022394:dJ511B24.2 (1-Phosphatidylinositol-4,5-Bisphosphate Phosphodiesterase Gamma 1 (EC 3.1.4.11, PLC-Gamma-1, Phospholipase C-Gamma-1 /cds=(68,3940) /gb=AL022394 /gi=3288442 /ug=Hs.317 /len=5151	34351_at
TRIP7 (thyroid hormone receptor interactor 7)	AA845349	Hs.77558		6	Cluster Incl. AA845349:ak01g01.s1 Homo sapiens cDNA, 3 end /clone=IMAGE-	37348_s_at

					1404720 /clone_end=3 /gb=AA845349 /gi=2933108 /ug=Hs.77558 /len=965	
YWHAQ (tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, theta polypeptide)	X56468	Hs.74405	NM_006826	22q12-qter	Cluster Incl. X56468:Human mRNA for 14.3.3 protein, a protein kinase regulator /cds=(125,862) /gb=X56468 /gi=23221 /ug=Hs.74405 /len=1862	32530_at
CGR19(cell growth regulatory with ring finger domain)	U66469	Hs.59106	NM_006568	14	U66469 /FEATURE= /DEFINITION=HSU66469 Human cell growth regulator CGR19 mRNA, complete cds	450_g_at
EZH1 (enhancer of zeste (Drosophila) homolog 1)	AB002386	Hs.194669	NM_001991	17q21.1-q21.3	Cluster Incl. AB002386:Human mRNA for KIAA0388 gene, complete cds /cds=(100,2343) /gb=AB002386 /gi=2224716 /ug=Hs.194669 /len=4606	32259_at
KIAA0746(KIAA0746 protein)	AB018289	Hs.49500		4	Cluster Incl. AB018289:Homo sapiens mRNA for KIAA0746 protein, partial cds /cds=(0,3091) /gb=AB018289 /gi=3882212	41585_at

					/ug=Hs.49500 /len=4086	
PRKCB1 (protein kinase C, beta 1)	X07109	Hs.77202	NM_002738	16p11.2	X07109 /FEATURE=cds /DEFINITION=HSPKCB2A Human mRNA for protein kinase C (PKC) type beta II	1217_g_at
TC21(oncogene TC21)	AI365215	Hs.206097	NM_012250	11	Cluster Incl. AI365215:qz41a06.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-2029426 /clone_end=3 /gb=AI365215 /gi=4124904 /ug=Hs.206097 /len=918	32827_at
RRN3(RNA polymerase I transcription factor RRN3)	AF001549	Hs.110103	NM_018427	16	Cluster Incl. AF001549:Human Chromosome 16 BAC clone CIT987SK-A-270G1 /cds=(167,487) /gb=AF001549 /gi=3355302 /ug=Hs.110103 /len=848	39820_at
XAP4(HBV associated factor)	AA160708	Hs.247280	NM_031229	20	Cluster Incl. AA160708:zo72c02.r1 Homo sapiens cDNA, 5 end /clone=IMAGE-592418 /clone_end=5 /gb=AA160708 /gi=1736075 /ug=Hs.18563 /len=643	32203_at

EEF2 (eukaryotic translation elongation factor 2)	Z11692	Hs.75309	NM_001961	19pter-q12	Cluster Incl. Z11692:H.sapiens mRNA for elongation factor 2 /cds=(0,2576) /gb=Z11692 /gi=31107 /ug=Hs.75309 /len=3080	36587_at
	U79277				Cluster Incl. U79277:Human clone 23548 mRNA sequence /cds=UNKNOWN /gb=U79277 /gi=1710245 /ug=Hs.71848 /len=1545	36760_at
KIAA0640(SWAP-70 protein)	AB014540	Hs.153026		11	Cluster Incl. AB014540:Homo sapiens mRNA for KIAA0640 protein, partial cds /cds=(0,1812) /gb=AB014540 /gi=3327093 /ug=Hs.153026 /len=4824	31869_at
BLNK (B-cell linker)	AF068180	Hs.167746	NM_013314	10q23.2-q23.33	Cluster Incl. AF068180:Homo sapiens B cell linker protein BLNK mRNA, alternatively spliced, complete cds /cds=(153,1523) /gb=AF068180 /gi=3406748 /ug=Hs.167746 /len=1790	38242_at

HNRPC (heterogeneous nuclear ribonucleoprotein C (C1/C2))	M16342	Hs.182447	NM_031314	2q32	Cluster Incl. M16342:Human nuclear ribonucleoprotein particle (hnRNP) C protein mRNA, complete cds /cds=UNKNOWN /gb=M16342 /gi=184266 /ug=Hs.182447 /len=1666	33666_at
KIAA0747(KIAA0747 protein)	AB018290	Hs.8309	NM_015292	12	Cluster Incl. AB018290:Homo sapiens mRNA for KIAA0747 protein, partial cds /cds=(0,3219) /gb=AB018290 /gi=3882214 /ug=Hs.8309 /len=4026	38424_at
M17S2 (membrane component, chromosome 17, surface marker 2 (ovarian carcinoma antigen CA125))	D30756	Hs.277721	NM_031858	17q21.1	Cluster Incl. D30756:Human mRNA for KIAA0049 gene, complete cds /cds=(140,3040) /gb=D30756 /gi=488500 /ug=Hs.233745 /len=4654	33444_at
CSK (c-src tyrosine kinase)	X59932	Hs.77793	NM_004383	15q23-q25	X59932 /FEATURE=mRNA /DEFINITION=HSCSRCKIN Human mRNA for C-SRC-kinase	1768_s_at
KIAA0239(KIAA0239 protein)	D87076	Hs.9729	NM_015288	5	Cluster Incl. D87076:Human mRNA for KIAA0239 gene, partial cds /cds=(0,1716)	38342_at

					/gb=D87076 /gi=1510152 /ug=Hs.9729 /len=5630	
NFATC1 (nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 1)	U08015	Hs.96149	NM_006162	18q23	Cluster Incl. U08015:Human NF-ATc mRNA, complete cds /cds=(239,2389) /gb=U08015 /gi=500631 /ug=Hs.96149 /len=2743	39143_at
TLK1 (tousled-like kinase 1)	D50927	Hs.18895	NM_012290	8p22-p12	Cluster Incl. D50927:Human mRNA for KIAA0137 gene, complete cds /cds=(1088,2737) /gb=D50927 /gi=1469196 /ug=Hs.18895 /len=4454	32219_at

Table 7:

UCL/HGNC/HUGO Human Gene Nomenclature Database Symbol	GenBank Accession No.	UniGene Cluster	RefSeq	Chromosomal Location	Description Unigene Build #95	Gene Name
MGC4175(hypothetical protein MGC4175)	AI656421	Hs.322404	NM_024315	7	Cluster Incl. AI656421:tt50h10.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-2244259 /clone_end=3 /gb=AI656421 /gi=4740400 /ug=Hs.5671 /len=566	41809_at
SH3BP5 (SH3-domain binding protein 5 (BTK-associated))	AB005047	Hs.109150	NM_004844	1q43	Cluster Incl. AB005047:Homo sapiens mRNA for SH3 binding protein, complete cds /cds=(63,1340) /gb=AB005047 /gi=3116213 /ug=Hs.109150 /len=2570	38968_at
MSF (MLL septin-like fusion (NOTE: non-standard symbol and name))	AB023208	Hs.181002	NM_006640	17q25	Cluster Incl. AB023208:Homo sapiens mRNA for KIAA0991 protein, complete cds /cds=(732,2000) /gb=AB023208 /gi=4589625 /ug=Hs.181002 /len=3938	41220_at

LEF1 (lymphoid enhancer-binding factor 1)	AL049409	Hs.44865	NM_016269	4q23-q25	Cluster Incl. AL049409:Homo sapiens mRNA; cDNA DKFZp586H0919 (from clone DKFZp586H0919) /cds=UNKNOWN /gb=AL049409 /gi=4500194 /ug=Hs.44865 /len=1419	36021_at
PRDM2 (PR domain containing 2, with ZNF domain)	D45132	Hs.26719	NM_012231	1p36	D45132 /FEATURE= /DEFINITION=HUMHOXY1 Homo sapiens mRNA for zinc-finger DNA-binding protein, complete cds	316_g_at
ARHH (ras homolog gene family, member H)	Z35227	Hs.109918	NM_004310	4p13	Cluster Incl. Z35227:H.sapiens TTF mRNA for small G protein /cds=(579,1154) /gb=Z35227 /gi=609016 /ug=Hs.109918 /len=1427	37416_at
SSH3BP1 (spectrin SH3 domain binding protein 1)	AF006516	Hs.24752	NM_005470	10p11.2	Cluster Incl. AF006516:Homo sapiens eps8 binding protein e3B1 mRNA, complete cds /cds=(66,1508) /gb=AF006516 /gi=2245670 /ug=Hs.24752 /len=3189	33886_at

ABLIM (actin binding LIM protein)	D31883	Hs.158203	NM_002313	10q25	Cluster Incl. D31883:Human mRNA for KIAA0059 gene, complete cds /cds=(221,1609) /gb=D31883 /gi=505093 /ug=Hs.158203 /len=6754	40155_at
SNRPA1 (small nuclear ribonucleoprotein polypeptide A')	X13482	Hs.80506	NM_003090	22q	Cluster Incl. X13482:Human mRNA for U2 snRNP-specific A protein /cds=(56,823) /gb=X13482 /gi=37546 /ug=Hs.80506 /len=1033	37585_at
UBE2D2 (ubiquitin-conjugating enzyme E2D 2 (homologous to yeast UBC4/5))	AI310002	Hs.108332	NM_003339	5p14.2-q23.3	Cluster Incl. AI310002:qo77c11.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-1914548 /clone_end=3 /gb=AI310002 /gi=4004873 /ug=Hs.108332 /len=656	38705_at
TERF1 (telomeric repeat binding factor (NIMA-interacting) 1)	U40705	Hs.194562	NM_017489	8q13	Cluster Incl. U40705:Homo sapiens telomeric repeat binding factor (TRF1) mRNA, complete cds /cds=UNKNOWN /gb=U40705 /gi=2078442 /ug=Hs.194562 /len=2686	32255_i_at

PPP3CC (protein phosphatase 3 (formerly 2B), catalytic subunit, gamma isoform (calcineurin A gamma))	S46622	Hs.75206	NM_005605	8	Cluster Incl. S46622:calcineurin A catalytic subunit [human, testis, mRNA, 2134 nt] /cds=(286,1794) /gb=S46622 /gi=258000 /ug=Hs.75206 /len=2134	32541_at
FIP2(tumor necrosis factor alpha-inducible cellular protein containing leucine zipper domains; Huntingtin interacting protein L; transcrption factor IIIA-interacting protein)	AF061034	Hs.278898	NM_021980	10	Cluster Incl. AF061034:Homo sapiens FIP2 alternatively translated mRNA, complete cds /cds=UNKNOWN /gb=AF061034 /gi=3127082 /ug=Hs.182236 /len=2116	41743_i_at
EP300 (E1A binding protein p300)	U01877	Hs.25272	NM_001429	22q13.2	Cluster Incl. U01877:Human p300 protein mRNA, complete cds /cds=(1199,8443) /gb=U01877 /gi=495300 /ug=Hs.25272 /len=9046	33896_at
PSCD1 (pleckstrin homology, Sec7 and coiled/coil domains 1(cytohesin 1))	M85169	Hs.1050	NM_004762	17q25	Cluster Incl. M85169:Human homologue of yeast sec7 mRNA, complete cds /cds=(69,1265) /gb=M85169 /gi=338001 /ug=Hs.1050 /len=3301	38666_at

SYNE-2(synaptic nuclei expressed gene 2)	AL080133	Hs.57749	NM_015180	14	Cluster Incl. AL080133:Homo sapiens mRNA; cDNA DKFZp434G173 (from clone DKFZp434G173) /cds=(122,3400) /gb=AL080133 /gi=5262573 /ug=Hs.57749 /len=4307	41815_at
BCL7A (B-cell CLL/lymphoma 7A)	X89984	Hs.211563	NM_020993	12q24.13	Cluster Incl. X89984:H.sapiens mRNA for BCL7A protein /cds=(953,1648) /gb=X89984 /gi=929614 /ug=Hs.211563 /len=4522	32842_at
RCN2 (reticulocalbin 2, EF-hand calcium binding domain)	X78669	Hs.79088	NM_002902	15q23	Cluster Incl. X78669:H.sapiens ERC-55 mRNA /cds=(66,1019) /gb=X78669 /gi=469884 /ug=Hs.79088 /len=1700	37727_i_at
CD79A (CD79A antigen (immunoglobulin-associated alpha))	U05259	Hs.79630	NM_001783	19q13.2	Cluster Incl. U05259:Human MB-1 gene, complete cds /cds=(36,716) /gb=U05259 /gi=452561 /ug=Hs.79630 /len=1107	38017_at
UBE2N (ubiquitin-conjugating enzyme E2N (homologous to yeast UBC13))	D83004	Hs.75355	NM_003348	12	Cluster Incl. D83004:Human epidermoid carcinoma mRNA for ubiquitin-conjugating enzyme E2 similar to Drosophila bendless gene product complete cds /cds=(63,521)	36604_at

					gene product, complete cds /cds=(63,521) /gb=D83004 /gi=1181557 /ug=Hs.75355 /len=1203	
DNTT (deoxynucleotidyltransferase, terminal)	M11722	Hs.272537	NM_004088	10q23-q24	Cluster Incl. M11722:Human terminal transferase mRNA, complete cds /cds=(328,1854) /gb=M11722 /gi=339436 /ug=Hs.234772 /len=2068	34168_at
WEE1 (wee1+ (S. pombe) homolog)	W28575	Hs.75188	NM_003390	11p15.3-p15.1	Cluster Incl. W28575:51f12 Homo sapiens cDNA /gb=W28575 /gi=1308730 /ug=Hs.8151 /len=906	38102_at
HSF2 (heat shock transcription factor 2)	Z99129	Hs.158195	NM_004506	6pter-p25.1	Cluster Incl. Z99129:Human DNA sequence from clone 425C14 on chromosome 6q22 Contains the HSF2 gene for Heat Shock Factor 2 (Heat Shock Transcription Factor 2, HSTF 2) and an unknown gene similar to the placental protein DIFF33 gene. Contains ESTs, STSs a	33443_at

ARFD1 (ADP-ribosylation factor domain protein 1, 64kD)	L04510	Hs.792	NM_001656	5	Cluster Incl. L04510:Human nucleotide binding protein mRNA, complete cds /cds=(22,1746) /gb=L04510 /gi=292069 /ug=Hs.792 /len=3334	37537_at
CTGF (connective tissue growth factor)	X78947	Hs.75511	NM_001901	6q23.1	Cluster Incl. X78947:H.sapiens mRNA for connective tissue growth factor /cds=(145,1194) /gb=X78947 /gi=474933 /ug=Hs.75511 /len=2312	36638_at
MALT1 (mucosa associated lymphoid tissue lymphoma translocation gene 1)	AB026118	Hs.180566	NM_006785	18q21	Cluster Incl. AB026118:Homo sapiens mRNA for MALT1, complete cds /cds=(65,2506) /gb=AB026118 /gi=5706377 /ug=Hs.188735 /len=2819	32350_at
SSH3BP1 (spectrin SH3 domain binding protein 1)	AF001628	Hs.24752	NM_005470	10p11.2	Cluster Incl. AF001628:Homo sapiens interactor protein AbIBP4 (AbIBP4) mRNA, complete cds /cds=(48,1403) /gb=AF001628 /gi=4100618 /ug=Hs.204036 /len=2175	38924_s_at

NR3C1 (nuclear receptor subfamily 3, group C, member 1)	M10901	Hs.75772	NM_000176	5q31	Cluster Incl. M10901:Human glucocorticoid receptor alpha mRNA, complete cds /cds=(132,2465) /gb=M10901 /gi=183032 /ug=Hs.75772 /len=4788	36690_at
TUBA1 (tubulin, alpha 1 (testis specific))	X06956	Hs.75318		2q	Cluster Incl. X06956:Human HALPHA44 gene for alpha-tubulin, exons 1-3 /cds=(0,1343) /gb=X06956 /gi=32014 /ug=Hs.75318 /len=1344	36591_at
KIAA0082(KIAA0082 protein)	D43949	Hs.154045		6	Cluster Incl. D43949:Human mRNA for KIAA0082 gene, partial cds /cds=(0,1824) /gb=D43949 /gi=603952 /ug=Hs.154045 /len=3186	40054_at
PRDM2 (PR domain containing 2, with ZNF domain)	D45132	Hs.26719	NM_012231	1p36	D45132 /FEATURE= /DEFINITION=HUMHOXY1 Homo sapiens mRNA for zinc-finger DNA-binding protein, complete cds	315_at

BARD1 (BRCA1 associated RING domain 1)	U76638	Hs.54089	NM_000465	2q34-q35	U76638 /FEATURE= /DEFINITION=HSU76638 Human BRCA1-associated RING domain protein (BARD1) mRNA, complete cds	1801_at
LYSAL1 (*lysosomal apyrase-like 1*)	AB002390	Hs.201377	NM_004901	8	Cluster Incl. AB002390:Human mRNA for KIAA0392 gene, partial cds /cds=(0,1652) /gb=AB002390 /gi=2280487 /ug=Hs.201377 /len=5422	33788_at
ARPP-19(cyclic AMP phosphoprotein, 19kD)	AL120559	Hs.7351	NM_006628	15	Cluster Incl. AL120559:DKFZp761B219_r1 Homo sapiens cDNA, 5 end /clone=DKFZp761B219 /clone_end=5 /gb=AL120559 /gi=5926458 /ug=Hs.7351 /len=588	36872_at
MADH7 (MAD (mothers against decapentaplegic, Drosophila) homolog 7)	AF010193	Hs.100602	NM_005904	18	AF010193 /FEATURE= /DEFINITION=AF010193 Homo sapiens MAD-related gene SMAD7 (SMAD7) mRNA, complete cds	1857_at

FLJ20500(hypothetical protein)	AA522530	Hs.111244	NM_019058	10	Cluster Incl. AA522530:ni38d12.s1 Homo sapiens cDNA, 3 end /clone=IMAGE-979127 /clone_end=3 /gb=AA522530 /gi=2263242 /ug=Hs.111244 /len=891	39827_at
					Glucocorticoid Receptor, Beta	706_at
	AF035315				Cluster Incl. AF035315:Homo sapiens clone 23664 and 23905 mRNA sequence /cds=UNKNOWN /gb=AF035315 /gi=2661077 /ug=Hs.180737 /len=1331	33267_at
SYNE-1B(synaptic nuclear envelope 1)	AB018339	Hs.8182		6	Cluster Incl. AB018339:Homo sapiens mRNA for KIAA0796 protein, partial cds /cds=(0,3243) /gb=AB018339 /gi=3882312 /ug=Hs.8182 /len=3900	38113_at
CTNNAL1 (catenin (cadherin-associated protein), alpha-like 1)	U97067	Hs.58488	NM_003798	9q31.2	Cluster Incl. U97067:Homo sapiens alpha-catenin-like protein mRNA, complete cds /cds=(43,2247) /gb=U97067 /gi=3342777 /ug=Hs.58488 /len=2446	35331_at

UBL3 (ubiquitin-like 3)	AL080177	Hs.173091	NM_007106	13q12-q13	Cluster Incl. AL080177:Homo sapiens mRNA; cDNA DKFZp434K151 (from clone DKFZp434K151) /cds=(109,462) /gb=AL080177 /gi=5262650 /ug=Hs.173091 /len=3313	40839_at
MICB (MHC class I polypeptide-related sequence B)	U65416	Hs.211580	NM_005931	6p21.3	Cluster Incl. U65416:Human MHC class I molecule (MICB) gene, complete cds /cds=(5,1156) /gb=U65416 /gi=1815636 /ug=Hs.211580 /len=2367	35937_at
DYRK3 (dual-specificity tyrosine-(Y)-phosphorylation regulated kinase 3)	Y12735	Hs.38018	NM_003582	1q32	Cluster Incl. Y12735:Homo sapiens mRNA for protein kinase, Dyrk3 /cds=(252,1913) /gb=Y12735 /gi=2765228 /ug=Hs.38018 /len=2141	39931_at
KIAA0256(KIAA0256 gene product)	D87445	Hs.118978		15	Cluster Incl. D87445:Human mRNA for KIAA0256 gene, complete cds /cds=(1424,3331) /gb=D87445 /gi=1665778 /ug=Hs.118978 /len=6935	41634_at

PPM1B (protein phosphatase 1B (formerly 2C), magnesium-dependent, beta isoform)	AJ005801	Hs.5687	NM_002706	2p21	Cluster Incl. AJ005801:Homo sapiens mRNA for protein phosphatase 2C (beta) /cds=(0,1439) /gb=AJ005801 /gi=3378167 /ug=Hs.169652 /len=1440	32665_at
FOXO1A (forkhead box O1A (rhabdomyosarcoma))	AF032885	Hs.170133	NM_002015	13q14.1	Cluster Incl. AF032885:Homo sapiens forkhead protein (FKHR) mRNA, complete cds /cds=(385,2352) /gb=AF032885 /gi=2895491 /ug=Hs.170133 /len=5723	40570_at
DUSP11 (dual specificity phosphatase 11 (RNA/RNP complex 1-interacting))	AF023917	Hs.14611	NM_003584	2	Cluster Incl. AF023917:Homo sapiens protein tyrosine phosphatase PIR1 mRNA, complete cds /cds=(124,1116) /gb=AF023917 /gi=3387789 /ug=Hs.14611 /len=1574	39727_at
FBXO21 (F-box only protein 21)	AB020682	Hs.184227		12	Cluster Incl. AB020682:Homo sapiens mRNA for KIAA0875 protein, partial cds /cds=(0,1866) /gb=AB020682 /gi=4240238 /ug=Hs.184227 /len=4168	32169_at

KIAA0105(Wilms' tumour 1-associating protein)	D14661	Hs.119	NM_004906	6	Cluster Incl. D14661:Human mRNA for KIAA0105 gene, complete cds /cds=(124,579) /gb=D14661 /gi=285946 /ug=Hs.119 /len=1622	41635_at
KIAA0922(KIAA0922 protein)	AB023139	Hs.37892	NM_015196	4	Cluster Incl. AB023139:Homo sapiens mRNA for KIAA0922 protein, partial cds /cds=(0,2372) /gb=AB023139 /gi=4589475 /ug=Hs.37892 /len=2505	39929_at
KIAA0118(KIAA0118 protein)	D42087	Hs.184627	NM_014999	12	Cluster Incl. D42087:Human mRNA for KIAA0118 gene, partial cds /cds=(0,485) /gb=D42087 /gi=576555 /ug=Hs.184627 /len=1413	33326_at
HNRPH2 (heterogeneous nuclear ribonucleoprotein H2 (H'))	U01923	Hs.278857	NM_019597	xq22	Cluster Incl. U01923:Human BTK region clone ftp-3 mRNA /cds=UNKNOWN /gb=U01923 /gi=460085 /ug=Hs.177025 /len=2220	41131_f_at
SCML2 (sex comb on midleg (Drosophila)-like 2)	Y18004	Hs.171558	NM_006089	xp22	Cluster Incl. Y18004:Homo sapiens mRNA for SCML2 protein /cds=(91,2193)	38518_at

2)					/gb=Y18004 /gi=4490941 /ug=Hs.171558 /len=4130	
MTMR6 (myotubularin related protein 6)	AF072928	Hs.79877		13q12	Cluster Incl. AF072928:Homo sapiens myotubularin related protein 6 mRNA, partial cds /cds=(0,1398) /gb=AF072928 /gi=3916215 /ug=Hs.79877 /len=3358	38035_at
RASGRP1 (RAS guanyl releasing protein 1 (calcium and DAG-regulated))	AF081195	Hs.182591	NM_005739	15q15	Cluster Incl. AF081195:Homo sapiens calcium and DAG-regulated guanine nucleotide exchange factor II mRNA, complete cds /cds=(103,2496) /gb=AF081195 /gi=3928854 /ug=Hs.182591 /len=5075	33291_at
IFITM1 (interferon induced transmembrane protein 1 (9-27))	J04164	Hs.146360	NM_003641	11	J04164 /FEATURE= /DEFINITION=HUM927A Human interferon-inducible protein 9-27 mRNA, complete cds	675_at
MAP-1(modulator of apoptosis 1)	AI670788	Hs.24719	NM_022151	14	Cluster Incl. AI670788:tz10c02.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-	34767_at

					2288162 /clone_end=3 /gb=AI670788 /gi=4850519 /ug=Hs.24719 /len=762	
SP3 (Sp3 transcription factor)	X68560	Hs.44450		2q31	Cluster Incl. X68560:H.sapiens SPR-2 mRNA for GT box binding protein /cds=(0,2094) /gb=X68560 /gi=38417 /ug=Hs.44450 /len=3504	41573_at
KLF7 (Kruppel-like factor 7 (ubiquitous))	AA478904	Hs.21599	NM_003709	2q32	Cluster Incl. AA478904:zv20c05.r1 Homo sapiens cDNA, 5 end /clone=IMAGE-754184 /clone_end=5 /gb=AA478904 /gi=2207538 /ug=Hs.21599 /len=577	34216_at
PIK3CA (phosphoinositide-3-kinase, catalytic, alpha polypeptide)	Z29090	Hs.85701	NM_006218	3q26.3	Cluster Incl. Z29090:H.sapiens mRNA for phosphatidylinositol 3-kinase /cds=(12,3218) /gb=Z29090 /gi=472990 /ug=Hs.85701 /len=3424	40704_at
ZNF161 (zinc finger protein 161)	D28118	Hs.6557	NM_007146	3q26.2	Cluster Incl. D28118:Human mRNA for DB1, complete cds /cds=(41,1591) /gb=D28118 /gi=529640 /ug=Hs.167558	32628_at

					/len=2306	
FIP2(tumor necrosis factor alpha-inducible cellular protein containing leucine zipper domains; Huntingtin interacting protein L; transcrption factor IIIA-interacting protein)	AF061034	Hs.278898	NM_021980	10	Cluster Incl. AF061034:Homo sapiens FIP2 alternatively translated mRNA, complete cds /cds=UNKNOWN /gb=AF061034 /gi=3127082 /ug=Hs.182236 /len=2116	41742_s_at
CDKN1B (cyclin-dependent kinase inhibitor 1B (p27, Kip1))	AI304854	Hs.238990	NM_004064	12p13.1-p12	Cluster Incl. AI304854:qo19f03.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-1908989 /clone_end=3 /gb=AI304854 /gi=3988543 /ug=Hs.238990 /len=625	33847_s_at
SMC4L1 (SMC4 (structural maintenance of chromosomes 4, yeast)-like 1)	AB019987	Hs.50758	NM_005496	3q26.1	Cluster Incl. AB019987:Homo sapiens mRNA for chromosome-associated polypeptide-C, complete cds /cds=(112,3978) /gb=AB019987 /gi=4092845 /ug=Hs.50758 /len=4086	34878_at
	AL050161				Cluster Incl. AL050161:Homo sapiens mRNA; cDNA DKFZp586B0222 (from clone DKFZp586B0222) /cds=UNKNOWN	40803_at

					/gb=AL050161 /gi=4884375 /ug=Hs.172089 /len=1573	
GMFB (glia maturation factor, beta)	AB001106	Hs.151413	NM_004124	14q22.1	AB001106 /FEATURE= /DEFINITION=AB001106 Homo sapiens mRNA for glia maturation factor, complete cds	763_at
CDKN1B (cyclin-dependent kinase inhibitor 1B (p27, Kip1))	AI304854	Hs.238990	NM_004064	12p13.1-p12	Cluster Incl. AI304854:qo19f03.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-1908989 /clone_end=3 /gb=AI304854 /gi=3988543 /ug=Hs.238990 /len=625	33848_r_at
TC21(oncogene TC21)	AI365215	Hs.206097	NM_012250	11	Cluster Incl. AI365215:qz41a06.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-2029426 /clone_end=3 /gb=AI365215 /gi=4124904 /ug=Hs.206097 /len=918	32827_at
KIAA1014(KIAA1014 protein)	AB023231	Hs.6834		13	Cluster Incl. AB023231:Homo sapiens mRNA for KIAA1014 protein, partial cds /cds=(0,2213) /gb=AB023231 /gi=4589677	35802_at

					/ug=Hs.6834 /len=3116	
RBM5 (RNA binding motif protein 5)	AF091263	Hs.201675	NM_005778	3p21.3	Cluster Incl. AF091263:Homo sapiens RNA binding motif protein 5 (RBM5) mRNA, complete cds /cds=(148,2595) /gb=AF091263 /gi=4140646 /ug=Hs.201675 /len=3097	32804_at
FLJ11220(hypothetical protein FLJ11220)	AL050064	Hs.3623	NM_018364	1	Cluster Incl. AL050064:Homo sapiens mRNA; cDNA DKFZp566L033 (from clone DKFZp566L033) /cds=UNKNOWN /gb=AL050064 /gi=4884294 /ug=Hs.129812 /len=2989	37828_at
SIAT9 (sialyltransferase 9 (CMP-NeuAc:lactosylceramide alpha-2,3-sialyltransferase; GM3 synthase))	AB018356	Hs.225939	NM_003896	2p24.3-p24.1	Cluster Incl. AB018356:Homo sapiens mRNA for GM3 synthase, complete cds /cds=(277,1365) /gb=AB018356 /gi=3779138 /ug=Hs.225939 /len=2359	34256_at
MTMR1 (myotubularin related protein 1)	AJ224979	Hs.23200		xq28	Cluster Incl. AJ224979:Homo sapiens mRNA for MTMR1 protein /cds=(0,1990) /gb=AJ224979 /gi=4128155 /ug=Hs.23200	34654_at

					/len=2582	
SPF30(splicing factor 30, survival of motor neuron-related)	AF107463	Hs.79968	NM_005871	10	Cluster Incl. AF107463:Homo sapiens splicing factor mRNA, complete cds /cds=(182,898) /gb=AF107463 /gi=3986747 /ug=Hs.79968 /len=1147	38040_at
ADNP (activity-dependent neuroprotector)	AB018327	Hs.3657	NM_015339	20q13.13-q13.2	Cluster Incl. AB018327:Homo sapiens mRNA for KIAA0784 protein, partial cds /cds=(0,3222) /gb=AB018327 /gi=3882288 /ug=Hs.3657 /len=4282	34394_at
KIAA0136(DNA segment, Chr 16, Johns Hopkins University 32, expressed)	D50926	Hs.70359		21q22.13	Cluster Incl. D50926:Human mRNA for KIAA0136 gene, partial cds /cds=(0,2854) /gb=D50926 /gi=1469194 /ug=Hs.70359 /len=4197	36845_at
SMG1(PI-3-kinase-related kinase SMG-1)	AC003007	Hs.110613	NM_014006	16	Cluster Incl. AC003007:Human Chromosome 16 BAC clone CIT987SK-A-61E3 /cds=(0,1742) /gb=AC003007 /gi=2911728 /ug=Hs.181634 /len=2410	41733_at

CCNE2 (cyclin E2)	AF091433	Hs.30464	NM_004702	8p22-q21.3	Cluster Incl. AF091433:Homo sapiens cyclin E2 mRNA, complete cds /cds=(37,1251) /gb=AF091433 /gi=3769613 /ug=Hs.30464 /len=2648	35249_at
TACC1 (transforming, acidic coiled-coil containing protein 1)	AF049910	Hs.173159	NM_006283	8p11	Cluster Incl. AF049910:Homo sapiens TACC1 (TACC1) mRNA, complete cds /cds=(320,2737) /gb=AF049910 /gi=3435156 /ug=Hs.173159 /len=7735	40841_at
RYBP (RING1 and YY1 binding protein)	AL049940	Hs.7910	NM_012234	3p21.1-cen	Cluster Incl. AL049940:Homo sapiens mRNA; cDNA DKFZp564E1922 (from clone DKFZp564E1922) /cds=UNKNOWN /gb=AL049940 /gi=4884183 /ug=Hs.7910 /len=3555	37732_at
TCF3 (transcription factor 3 (E2A immunoglobulin enhancer binding factors E12/E47))	M31523	Hs.101047		19p13.3	M31523 /FEATURE= /DEFINITION=HUMTFAA Human transcription factor (E2A) mRNA, complete cds	1373_at

KLRB1 (killer cell lectin-like receptor subfamily B, member 1)	U11276	Hs.169824	NM_002258	12p13	Cluster Incl. U11276:Human hNKR-P1a protein (NKR-P1A) mRNA, complete cds /cds=(60,737) /gb=U11276 /gi=538270 /ug=Hs.169824 /len=738	35449_at
TERF1 (telomeric repeat binding factor (NIMA-interacting) 1)	U74382	Hs.194562	NM_017489	8q13	U74382 /FEATURE= /DEFINITION=HSU74382 Human telomeric repeat DNA-binding protein (PIN2) mRNA, complete cds	1329_s_at
BRD1 (bromodomain-containing 1)	Z98885	Hs.127950	NM_014577	22q13.33	Cluster Incl. Z98885:Human DNA sequence from clone 522J7 on chromosome 22q13.3. Contains part of a 60S Ribosomal protein L5 pseudogene and a Peregrin (BR140) LIKE gene downstream of a putative CpG island. Contains ESTs, STSs and GSSs /cds=(185,3361) /gb=Z	39894_f_at
TGFBR2 (transforming growth factor, beta receptor II (70-80kD))	D50683	Hs.82028	NM_003242	3p22	D50683 /FEATURE= /DEFINITION=D50683 Homo sapiens mRNA for TGF-betaIIR alpha, complete	1815_g_at

					cds	
LPIN1 (lipin 1)	D80010	Hs.81412		2p21	Cluster Incl. D80010:Human mRNA for KIAA0188 gene, partial cds /cds=(0,2700) /gb=D80010 /gi=1136435 /ug=Hs.81412 /len=5307	38098_at
DKFZP586J0619(DKFZP586J0619 protein)	AL050110	Hs.112184			Cluster Incl. AL050110:Homo sapiens mRNA; cDNA DKFZp586J0619 (from clone DKFZp586J0619) /cds=(0,1923) /gb=AL050110 /gi=4884139 /ug=Hs.112184 /len=2224	39022_at
CASP7 (caspase 7, apoptosis-related cysteine protease)	U67319	Hs.9216	NM_001227	10q25	Cluster Incl. U67319:Human Lice2 beta cysteine protease mRNA, complete cds /cds=(228,1238) /gb=U67319 /gi=1894912 /ug=Hs.9216 /len=2602	38281_at
DCK (deoxycytidine kinase)	M60527	Hs.709	NM_000788	4q13.3-q21.1	M60527 /FEATURE=mRNA /DEFINITION=HUMDCKATPB Human deoxycytidine kinase mRNA, complete cds	886_at

KIAA0080(KIAA0080 protein)	D38522	Hs.74554		1	Cluster Incl. D38522:Human mRNA for KIAA0080 gene, partial cds /cds=(0,318) /gb=D38522 /gi=559331 /ug=Hs.74554 /len=4001	36144_at
POU2AF1 (POU domain, class 2, associating factor 1)	Z49194	Hs.2407	NM_006235	11q23.1	Cluster Incl. Z49194:H.sapiens mRNA for oct-binding factor /cds=(523,1293) /gb=Z49194 /gi=974830 /ug=Hs.2407 /len=3301	36239_at
LOC57158(hypothetical protein LOC57158)	AL035447	Hs.134594	NM_020433	20	Cluster Incl. AL035447:Human DNA sequence from clone 1183I21 on chromosome 20q12. Contains a novel gene and the first exon of a putative novel gene for a protein similar to predicted fly and worm proteins. Contains ESTs, STSs, GSSs and a putative CpG isla	36934_at
KIAA0794(KIAA0794 protein)	AB018337	Hs.127287		3	Cluster Incl. AB018337:Homo sapiens mRNA for KIAA0794 protein, partial cds /cds=(0,1472) /gb=AB018337 /gi=3882308	41691_at

					/ug=Hs.127287 /len=4656	
CD8A (CD8 antigen, alpha polypeptide (p32))	M12824	Hs.85258	NM_001768	2p12	Cluster Incl. M12824:Human T-cell differentiation antigen Leu-2/T8 mRNA, partial cds /cds=(87,794) /gb=M12824 /gi=339426 /ug=Hs.85258 /len=1975	40699_at
KIAA0275(KIAA0275 gene product)	D87465	Hs.74583	NM_014767	10	Cluster Incl. D87465:Human mRNA for KIAA0275 gene, complete cds /cds=(316,1590) /gb=D87465 /gi=1665814 /ug=Hs.74583 /len=5316	36155_at
FLJ20259(hypothetical protein FLJ20259)	W27545	Hs.9956	NM_017730		Cluster Incl. W27545:32c4 Homo sapiens cDNA /gb=W27545 /gi=1307349 /ug=Hs.9956 /len=950	38362_at
EP300 (E1A binding protein p300)	U01877	Hs.25272	NM_001429	22q13.2	U01877 /FEATURE= /DEFINITION=HSU01877 Human p300 protein mRNA, complete cds	551_at

SEC24B (SEC24 (S. cerevisiae) related gene family, member B)	AJ131245	Hs.7239	NM_006323	4	Cluster Incl. AJ131245:Homo sapiens mRNA for Sec24 protein (Sec24B isoform) /cds=(155,3961) /gb=AJ131245 /gi=3947689 /ug=Hs.7239 /len=4742	35845_at
SCYE1 (small inducible cytokine subfamily E, member 1 (endothelial monocyte-activating))	U10117	Hs.333513	NM_004757	4q24-q26	Cluster Incl. U10117:Human endothelial-monocyte activating polypeptide II mRNA, complete cds /cds=(49,987) /gb=U10117 /gi=498909 /ug=Hs.146401 /len=1057	39734_at
MADH3 (MAD (mothers against decapentaplegic, Drosophila) homolog 3)	U68019	Hs.211578	NM_005902	15q21-q22	U68019 /FEATURE= /DEFINITION=HSU68019 Homo sapiens mad protein homolog (hMAD-3) mRNA, complete cds	1433_g_at
OS-9(amplified in osteosarcoma)	U41635	Hs.76228	NM_006812	12	Cluster Incl. U41635:Human OS-9 precurosor mRNA, complete cds /cds=(85,2088) /gb=U41635 /gi=1322233 /ug=Hs.76228 /len=2736	36996_at

PLOD3 (procollagen-lysine, 2-oxoglutarate 5-dioxygenase 3)	AF046889	Hs.153357	NM_001084	7q22	Cluster Incl. AF046889:Homo sapiens lysyl hydroxylase isoform 3 (PLOD3) mRNA, complete cds /cds=(216,2432) /gb=AF046889 /gi=3153234 /ug=Hs.153357 /len=2735	39801_at
APLP2 (amyloid beta (A4) precursor-like protein 2)	S60099	Hs.279518	NM_001642	11q24	Cluster Incl. S60099:APPH=amyloid precursor protein homolog [human, placenta, mRNA, 3727 nt] /cds=(72,2363) /gb=S60099 /gi=300168 /ug=Hs.64797 /len=3727	33944_at
NME2 (non-metastatic cells 2, protein (NM23B) expressed in)	X58965	Hs.275163	NM_002512	17q21.3	X58965 /FEATURE= /DEFINITION=HSNM23H2G H.sapiens RNA for nm23-H2 gene	1980_s_at
MPB1 (MYC promoter-binding protein 1)	M55914	Hs.284127	NM_005945	1pter-p35	M55914 /FEATURE= /DEFINITION=HUMCMYCQ Human c-myc binding protein (MBP-1) mRNA, complete cds	2035_s_at

CST3 (cystatin C (amyloid angiopathy and cerebral hemorrhage))	AI362017	Hs.135084	NM_000099	20p11.2	Cluster Incl. AI362017:qy39a10.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-2014362 /clone_end=3 /gb=AI362017 /gi=4113638 /ug=Hs.135084 /len=778	39689_at
RPN2 (ribophorin II)	AL031659	Hs.75722	NM_002951	20q12-q13.1	Cluster Incl. AL031659:dJ343K2.2.1 (ribophorin II (isoform 1)) /cds=(284,2179) /gb=AL031659 /gi=4468296 /ug=Hs.75722 /len=2488	36676_at

Table 8:

UCL/HGNC/HUGO Human Gene Nomenclature Database Symbol	GenBank Accession No.	UniGene Cluster	RefSeq	Chromosomal Location	Description Unigene Build #95	Gene Name
SGP28(specific granule protein (28 kDa); cysteine-rich secretory protein-3	X94323	Hs.54431	NM_006061	6	Cluster Incl. X94323:H.sapiens mRNA for SGP28 protein /cds=(40,777) /gb=X94323 /gi=1213612 /ug=Hs.54431 /len=2124	36464_at
MMP9 (matrix metalloproteinase 9 (gelatinase B, 92kD gelatinase, 92kD type IV collagenase)	J05070	Hs.151738	NM_004994	20q11.2-q13.1	Cluster Incl. J05070:Human type IV collagenase mRNA, complete cds /cds=(19,2142) /gb=J05070 /gi=177204 /ug=Hs.151738 /len=2334	31859_at
FLJ10140(hypothetical protein FLJ10140)	AL031588	Hs.250671	NM_018006	22	Cluster Incl. AL031588:dJ1163J1.3 (novel protein similar to mouse B99) /cds=(0,2140) /gb=AL031588 /gi=4007108 /ug=Hs.122552 /len=2821	39872_at

CAMP (cathelicidin antimicrobial peptide)	Z38026	Hs.51120	NM_004345	3p21.3	Cluster Incl. Z38026:H.sapiens mRNA for FALL-39 peptide antibiotic /cds=(11,523) /gb=Z38026 /gi=558378 /ug=Hs.51120 /len=615	36710_at
LCN2 (lipocalin 2 (oncogene 24p3)	AI762213	Hs.204238	NM_005564	9q34	Cluster Incl. AI762213:wi54d04.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-2394055 /clone_end=3 /gb=AI762213 /gi=5177880 /ug=Hs.204238 /len=677	32821_at
UGCG (UDP-glucose ceramide glucosyltransferase	D50840	Hs.152601	NM_003358	9q31	Cluster Incl. D50840:Homo sapiens mRNA for ceramide glucosyltransferase, complete cds /cds=(290,1474) /gb=D50840 /gi=1350551 /ug=Hs.152601 /len=1637	40215_at
KLF5 (Kruppel-like factor 5 (intestinal)	D14520	Hs.84728	NM_001730	13q21.2-13q22.2	Cluster Incl. D14520:Human mRNA for GC-Box binding protein BTEB2, complete cds /cds=(558,1217) /gb=D14520 /gi=303596 /ug=Hs.84728 /len=1301	37926_at

SCYC2 (small inducible cytokine subfamily C, member 2)	D63789	Hs.174228	NM_003175	1q23-q25	Cluster Incl. D63789:Homo sapiens DNA for SCM-1beta precursor, complete cds /cds=(21,365) /gb=D63789 /gi=1754608 /ug=Hs.174228 /len=485	31495_at
DEFA4 (defensin, alpha 4, corticostatin)	AI250799	Hs.2582	NM_001925	8p23	Cluster Incl. AI250799:qi36g07.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-1858620 /clone_end=3 /gb=AI250799 /gi=3847328 /ug=Hs.2582 /len=542	34546_at
SYNE-1B(synaptic nuclear envelope 1)	AB018339	Hs.8182		6	Cluster Incl. AB018339:Homo sapiens mRNA for KIAA0796 protein, partial cds /cds=(0,3243) /gb=AB018339 /gi=3882312 /ug=Hs.8182 /len=3900	38113_at
CCR2 (chemokine (C-C motif) receptor 2)	U95626	Hs.395	NM_000647	3p21	Cluster Incl. U95626:Homo sapiens ccr2b (ccr2), ccr2a (ccr2), ccr5 (ccr5) and ccr6 (ccr6) genes, complete cds, and lactoferrin (lactoferrin) gene, partial cds /cds=(2,1429) /gb=U95626 /gi=2104517 /ug=Hs.105938 /len=1607	37149_s_at

CLC (Charot-Leyden crystal protein)	L01664	Hs.132004	NM_013246	11q13.3	Cluster Incl. L01664:Human eosinophil Charcot-Leyden crystal (CLC) protein (lysophospholipase) mRNA, complete cds /cds=(33,461) /gb=L01664 /gi=187273 /ug=Hs.889 /len=586	36809_at
CEACAM8 (carcinoembryonic antigen-related cell adhesion molecule 8)	M33326	Hs.41	NM_001816	19q13.2	Cluster Incl. M33326:Human nonspecific cross-reacting antigen - (NCA) mRNA, complete cds /cds=(86,1135) /gb=M33326 /gi=189101 /ug=Hs.41 /len=2287	33530_at
CYP4F3 (cytochrome P450, subfamily IVF, polypeptide 3 (leukotriene B4 omega hydroxylase))	D12620	Hs.106242	NM_000896	19p13.2	D12620 /FEATURE= /DEFINITION=HUMCYT1 Homo sapiens mRNA for cytochrome P-450LTBV, complete cds	1305_s_at
KIAA0601(KIAA0601 protein)	W28504	Hs.174174		1	Cluster Incl. W28504:48e7 Homo sapiens cDNA /gb=W28504 /gi=1308515 /ug=Hs.154085 /len=1007	36338_at
LOC96807(hypothetical gene supported by X89214; NM_020995	X89214			16	Cluster Incl. X89214:H.sapiens mRNA for haptoglobin related protein	36984_f_at

X89214; NM_020995					/cds=(138,1295) /gb=X89214 /gi=1495457 /ug=Hs.75990 /len=1460	
CEACAM6 (carcinoembryonic antigen-related cell adhesion molecule 6 (non-specific cross	M18728	Hs.73848	NM_002483	19q13.2	Cluster Incl. M18728:Human nonspecific crossreacting antigen mRNA, complete cds /cds=UNKNOWN /gb=M18728 /gi=189084 /ug=Hs.73848 /len=2533	36105_at
CDA (cytidine deaminase)	L27943	Hs.72924	NM_001785	1p36.2-p35	L27943 /FEATURE=mRNA /DEFINITION=HUMCYDE Homo sapiens cytidine deaminase (CDA) mRNA, complete cds	1117_at
ARG1 (arginase, liver)	M14502	Hs.289057	NM_000045	6q23	M14502 /FEATURE=mRNA /DEFINITION=HUMARGL Human liver arginase mRNA, complete cds	1962_at
BPI (bactericidal/permeability-increasing protein)	J04739	Hs.89535	NM_001725	20q11.23-q12	Cluster Incl. J04739:Human bactericidal permeability increasing protein (BPI) mRNA, complete cds /cds=(30,1493) /gb=J04739 /gi=179528 /ug=Hs.89535	37054_at

					/len=1813	
MMP8 (matrix metalloproteinase 8 (neutrophil collagenase)	J05556	Hs.73862	NM_002424	11q22.3	J05556 /FEATURE=mRNA /DEFINITION=HUMCLGNA Homo sapiens collagenase mRNA, complete cds	681_at
BN51T (BN51 (BHK21) temperature sensitivity complementing	M17754	Hs.1276	NM_001722	8q21	Cluster Incl. M17754:Human BN51 mRNA, complete cds /cds=(51,1238) /gb=M17754 /gi=179512 /ug=Hs.1276 /len=1881	41694_at
MME (membrane metallo-endopeptidase (neutral endopeptidase, enkephalinase, CALLA, CD10)	J03779	Hs.1298	NM_000902	3q25.1-q25.2	J03779 /FEATURE=mRNA /DEFINITION=HUMCALLA Human common acute lymphoblastic leukemia antigen (CALLA) mRNA, complete cds	1389_at
RBM9 (RNA binding motif protein 9	AL009266	Hs.5011	NM_014309	22q13.1	Cluster Incl. AL009266:H. sapiens cDNA similar to C. elegans RNA binding protein U14946, Q10572, complete cds /cds=(170,1273) /gb=AL009266 /gi=2664428 /ug=Hs.155156 /len=1876	40260_g_at

LBP (lipopolysaccharide-binding protein	AF013512	Hs.154078	NM_004139	20q11.23-q12	Cluster Incl. AF013512:untitled /cds=(106,1551) /gb=AF013512 /gi=2653816 /ug=Hs.154078 /len=1887	35013_at
ORM1 (orosomucoid 1)	X02544	Hs.572	NM_000607	9q31-q32	Cluster Incl. X02544:Human mRNA for alpha1-acid glycoprotein (orosomucoid) /cds=(78,683) /gb=X02544 /gi=24444 /ug=Hs.572 /len=803	35315_at
CD24 (CD24 antigen (small cell lung carcinoma cluster 4 antigen)	L33930	Hs.286124	NM_013230	6q21	L33930 /FEATURE= /DEFINITION=HUMCD24B Homo sapiens CD24 signal transducer mRNA, complete cds and 3 region	266_s_at
NCF4 (neutrophil cytosolic factor 4 (40kD)	AL008637	Hs.196352	NM_000631	22q13.1	Cluster Incl. AL008637:Human DNA sequence from clone 833B7 on chromosome 22q12.3-13.2 Contains genes for NCF4 (P40PHOX) protein,cytokine receptor common beta chain precursor CSF2RB (partial), ESTs, CA repeat, STS, GSS /cds=(629,1648)	38894_g_at

					/gb=AL008637 /gi=3136	
	AF035315				Cluster Incl. AF035315:Homo sapiens clone 23664 and 23905 mRNA sequence /cds=UNKNOWN /gb=AF035315 /gi=2661077 /ug=Hs.180737 /len=1331	33267_at
CEACAM1 (carcinoembryonic antigen-related cell adhesion molecule 1 (biliary glycoprotein	X16354	Hs.50964	NM_001712	19q13.2	X16354 /FEATURE= /DEFINITION=HSTM1CEA Human mRNA for transmembrane carcinoembryonic antigen BGPa (formerly TM1-CEA)	988_at
QPCT (glutaminyl-peptide cyclotransferase (glutaminyl cyclase	X71125	Hs.79033	NM_012413	2p22.3-2p22.1	Cluster Incl. X71125:H.sapiens mRNA for glutamine cyclotransferase /cds=UNKNOWN /gb=X71125 /gi=398375 /ug=Hs.234747 /len=1556	35966_at

DEFA1 (defensin, alpha 1, myeloid-related sequence	AL036554	Hs.274463	NM_004084	8p23.2-p23.1	Cluster Incl. AL036554:DKFZp564J2262_r1 Homo sapiens cDNA, 5 end /clone=DKFZp564J2262 /clone_end=5 /gb=AL036554 /gi=5927801 /ug=Hs.1379 /len=517	31793_at
NR2F6 (nuclear receptor subfamily 2, group F, member 6)	AI189624	Hs.239752	NM_005234	19p13.1	Cluster Incl. AI189624:qd32h08.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-1725471 /clone_end=3 /gb=AI189624 /gi=3740833 /ug=Hs.239752 /len=833	41312_r_at
TCN1 (transcobalamin I (*vitamin B12 binding protein*, R binder family)	J05068	Hs.2012	NM_001062	11q11-q12	Cluster Incl. J05068:human transcobalamin I mRNA, complete cds /cds=(75,1376) /gb=J05068 /gi=307478 /ug=Hs.2012 /len=1537	35919_at
IL8RA (interleukin 8 receptor, alpha)	U11870	Hs.194778	NM_000634	2q35	U11870 /FEATURE=mRNA /DEFINITION=HSU11870 Human interleukin-8 receptor type A (IL8RBA) gene, promoter and complete cds	1352_at

KIAA0604(KIAA0604 gene product	AB011176	Hs.129801	NM_014693	3	Cluster Incl. AB011176:Homo sapiens mRNA for KIAA0604 protein, complete cds /cds=(138,2435) /gb=AB011176 /gi=3043731 /ug=Hs.129801 /len=3228	35536_at
FRAT2 (*frequently rearranged in advanced T-cell lymphomas 2*	AF062739	Hs.140720		10q23-q24.1	Cluster Incl. AF062739:Homo sapiens GSK-3 binding protein FRAT2 (FRAT2) mRNA, partial cds /cds=(0,341) /gb=AF062739 /gi=3243176 /ug=Hs.140720 /len=481	40171_at
LOC90355(hypothetical gene supported by AF038182; BC009203	AF038182	Hs.25925		5	Cluster Incl. AF038182:Homo sapiens clone 23860 mRNA sequence /cds=UNKNOWN /gb=AF038182 /gi=2795902 /ug=Hs.25925 /len=1508	33466_at
CPNE3 (copine III)	AB014536	Hs.14158	NM_003909	8p22-q21.3	Cluster Incl. AB014536:Homo sapiens mRNA for KIAA0636 protein, complete cds /cds=(120,1733) /gb=AB014536 /gi=3327085 /ug=Hs.14158 /len=4737	39706_at

GAS11 (growth arrest-specific 11	AF050078	Hs.54877	NM_001481	16q24.3	Cluster Incl. AF050078:untitled /cds=(122,1558) /gb=AF050078 /gi=3818466 /ug=Hs.54877 /len=3186	36479_at
CHIT1 (chitinase 1 (chitotriosidase)	U29615	Hs.91093	NM_003465	1q31-q32	Cluster Incl. :Human chitotriosidase precursor mRNA, complete cds /cds=(12,1412) /gb=U29615 /gi=1050957 /ug=Hs.91093 /len=1633	37061_at
RAB31 (RAB31, member RAS oncogene family	U59877	Hs.223025	NM_006868	18p11.3	Cluster Incl. U59877:Human low-Mr GTP-binding protein (RAB31) mRNA, complete cds /cds=(60,644) /gb=U59877 /gi=1388194 /ug=Hs.223025 /len=907	33371_s_at
TUBA1 (tubulin, alpha 1 (testis specific)	X06956	Hs.75318		2q3	Cluster Incl. X06956:Human HALPHA44 gene for alpha-tubulin, exons 1-3 /cds=(0,1343) /gb=X06956 /gi=32014 /ug=Hs.75318 /len=1344	36591_at
STX3A (syntaxin 3A	U32315	Hs.82240	NM_004177	11cen-11q12.3	Cluster Incl. U32315:Human syntaxin 3 mRNA, complete cds /cds=(38,907) /gb=U32315 /gi=929990 /ug=Hs.82240	38381_at

					/len=1903	
	AA524802				Cluster Incl. AA524802:nh33h11.s1 Homo sapiens cDNA /clone=IMAGE-954213 /gb=AA524802 /gi=2265730 /ug=Hs.203907 /len=500	32877_i_at
PDI2(peptidyl arginine deiminase, type II)	AB023211	Hs.33455		1	Cluster Incl. AB023211:Homo sapiens mRNA for KIAA0994 protein, partial cds /cds=(0,2061) /gb=AB023211 /gi=4589631 /ug=Hs.33455 /len=4343	35674_at
ITGAM (integrin, alpha M (complement component receptor 3, alpha; also known as CD11b (p170),	J03925	Hs.172631	NM_000632	16p11.2	Cluster Incl. J03925:Human Mac-1 gene encoding complement receptor type 3, CD11b, complete cds /cds=(72,3533) /gb=J03925 /gi=187284 /ug=Hs.172631 /len=4699	38533_s_at
ALOX5AP (arachidonate 5-lipoxygenase-activating protein	AI806222	Hs.100194	NM_001629	13q12	Cluster Incl. AI806222:wf26e10.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-2356746 /clone_end=3 /gb=AI806222	37099_at

					/gi=5392788 /ug=Hs.100194 /len=563	
DEFA3 (defensin, alpha 3, neutrophil-specific	L12691	Hs.294176	NM_005217	8pter-p23.3	Cluster Incl. L12691:Human neutrophil peptide-3 gene, complete cds /cds=(50,334) /gb=L12691 /gi=292364 /ug=Hs.178741 /len=452	31506_s_at
P63(transmembrane protein (63kD), endoplasmic reticulum/Golgi intermediate compartment	X69910	Hs.74368	NM_006825	12	Cluster Incl. X69910:H.sapiens p63 mRNA for transmembrane protein /cds=(84,1889) /gb=X69910 /gi=297407 /ug=Hs.74368 /len=2898	32529_at
ANXA3 (annexin A3	M20560	Hs.1378	NM_005139	4q13-q22	Cluster Incl. M20560:Human lipocortin-III mRNA, complete cds /cds=(46,1017) /gb=M20560 /gi=186967 /ug=Hs.1378 /len=1339	31792_at
					Tubulin, Alpha 1, Isoform 44	330_s_at
SLC2A5 (solute carrier family 2 (facilitated glucose/fructose transporter), member 5	M55531	Hs.33084	NM_003039	1p36.2	Cluster Incl. M55531:Human glucose transport-like 5 (GLUT5) mRNA, complete cds /cds=(75,1580) /gb=M55531	34362_at

					/gi=183297 /ug=Hs.33084 /len=2218	
SIM2 (single-minded (Drosophila) homolog 2	U80457	Hs.27311	NM_005069	21q22.13	Cluster Incl. U80457:Human transcription factor SIM2 short form mRNA, complete cds /cds=(92,1804) /gb=U80457 /gi=2062418 /ug=Hs.27311 /len=2844	39609_at
	U72507				Cluster Incl. U72507:Human 40871 mRNA partial sequence /cds=UNKNOWN /gb=U72507 /gi=1673508 /ug=Hs.234216 /len=1414	39245_at
LCP2 (lymphocyte cytosolic protein 2 (SH2 domain-containing leukocyte protein of 76kD	U20158	Hs.2488	NM_005565	5q33.1-qter	Cluster Incl. U20158:Human 76 kDa tyrosine phosphoprotein SLP-76 mRNA, complete cds /cds=(255,1656) /gb=U20158 /gi=806765 /ug=Hs.2488 /len=2032	39319_at

NCF4 (neutrophil cytosolic factor 4 (40kD)	AL008637	Hs.196352	NM_000631	22q13.1	Cluster Incl. AL008637:Human DNA sequence from clone 833B7 on chromosome 22q12.3-13.2 Contains genes for NCF4 (P40PHOX) protein,cytokine receptor common beta chain precursor CSF2RB (partial), ESTs, CA repeat, STS, GSS /cds=(629,1648) /gb=AL008637 /gi=3136	38893_at
MYL2 (myosin, light polypeptide 2, regulatory, cardiac, slow	X66141	Hs.75535	NM_000432	12q23-q24.3	Cluster Incl. X66141:H.sapiens mRNA for cardiac ventricular myosin light chain-2 /cds=(30,530) /gb=X66141 /gi=34845 /ug=Hs.75535 /len=784	36640_at
PPP2R5A (protein phosphatase 2, regulatory subunit B (B56), alpha isoform	L42373	Hs.155079	NM_006243	1q41	L42373 /FEATURE=mRNA /DEFINITION=HUMPP2A Homo sapiens phosphatase 2A B56-alpha (PP2A) mRNA, complete cds	903_at
FCN1 (ficolin (collagen/fibrinogen domain-containing) 1	S80990	Hs.252136	NM_002003	9q34	Cluster Incl. S80990:ficolin [human, uterus, mRNA, 1736 nt] /cds=(532,1512) /gb=S80990 /gi=1911529 /ug=Hs.169237	36447_at

					/len=1723	
IL8RA (interleukin 8 receptor, alpha	U11870	Hs.194778	NM_000634	2q35	U11870 /FEATURE=mRNA /DEFINITION=HSU11870 Human interleukin-8 receptor type A (IL8RBA) gene, promoter and complete cds	1353_g_at
NFIB (nuclear factor I/B	U85193	Hs.33287	NM_005596	9p24.1	Cluster Incl. :Human nuclear factor I-B2 (NFIB2) mRNA, complete cds /cds=(209,1471) /gb=U85193 /gi=1814408 /ug=Hs.239235 /len=2424	34720_at
NS1-BP(NS1-binding protein)	AB020657	Hs.197298	NM_006469	1	Cluster Incl. AB020657:Homo sapiens mRNA for KIAA0850 protein, complete cds /cds=(630,2558) /gb=AB020657 /gi=4240188 /ug=Hs.197298 /len=3682	33752_at
PLXNC1 (plexin C1	AF030339	Hs.286229	NM_005761	12	Cluster Incl. AF030339:Homo sapiens receptor for viral semaphorin protein (VESPR) mRNA, complete cds /cds=(249,4955) /gb=AF030339	32193_at

					/gi=3176761 /ug=Hs.184697 /len=5121	
AIF1 (allograft inflammatory factor 1	Y14768	Hs.76364	NM_001623	6p21.3	Cluster Incl. Y14768:Homo sapiens DNA, cosmid clones TN62 and TN82 /cds=(10,744) /gb=Y14768 /gi=3805800 /ug=Hs.890 /len=896	40729_s_at
CD59 (CD59 antigen p18-20 (antigen identified by monoclonal antibodies 16.3A5, EJ16, EJ30, EL32	M84349	Hs.119663	NM_000611	11p13	Cluster Incl. M84349:Human transmembrane protein (CD59) gene /cds=(18,404) /gb=M84349 /gi=180150 /ug=Hs.119663 /len=1840	39351_at
SLPI (secretory leukocyte protease inhibitor (antileukoproteinase)	X04470	Hs.251754	NM_003064	20pter-p12.3	Cluster Incl. X04470:Human mRNA for antileukoprotease (ALP) from cervix uterus /cds=(18,416) /gb=X04470 /gi=28638 /ug=Hs.169793 /len=594	32275_at
OGG1 (8-oxoguanine DNA glycosylase	AB019529	Hs.96398	NM_002542	3p26.2	Cluster Incl. AB019529:Homo sapiens mRNA for OGG1 protein type 2c, partial cds /cds=(0,303) /gb=AB019529	34146_at

					/gi=4587151 /ug=Hs.227236 /len=585	
CEACAM1 (carcinoembryonic antigen-related cell adhesion molecule 1 (biliary glycoprotein)	S71326	Hs.50964	NM_001712	19q13.2	Cluster Incl. S71326:BGPc=biliary glycoprotein adhesion molecule {alternatively spliced} [human, HT29 colon carcinoma cell line, mRNA Partial, 1473 nt] /cds=(0,1394) /gb=S71326 /gi=550030 /ug=Hs.50964 /len=1473	36082_at
KIAA0374(syntaphilin)	AB002372			20	Cluster Incl. AB002372:Human mRNA for KIAA0374 gene, complete cds /cds=(642,2258) /gb=AB002372 /gi=2224688 /ug=Hs.100837 /len=5530	41107_at
KIAA1564(KIAA1564 protein	U00930	Hs.173421		14	Cluster Incl. U00930:Human clone C4E 1.63 (CAC)n/(GTG)n repeat-containing mRNA /cds=UNKNOWN /gb=U00930 /gi=405043 /ug=Hs.204196 /len=3276	40981_at
LTA4H (leukotriene A4 hydrolase)	J03459	Hs.81118	NM_000895	12q22	Cluster Incl. J03459:Human leukotriene A-4 hydrolase mRNA, complete cds /cds=(68,1903) /gb=J03459 /gi=187172	38081_at

					/ug=Hs.81118 /len=2060	
COL17A1 (collagen, type XVII, alpha 1)	M91669	Hs.117938	NM_000494	10q24.3	Cluster Incl. M91669:Human Bullous pemphigoid autoantigen BP180 gene, 3 end /cds=(0,4598) /gb=M91669 /gi=179516 /ug=Hs.117938 /len=4669	41618_at
EPB72 (erythrocyte membrane protein band 7.2 (stomatin)	X85116	Hs.160483	NM_004099	9q34.1	Cluster Incl. X85116:H.sapiens epb72 gene exon 1 /cds=(61,927) /gb=X85116 /gi=1161561 /ug=Hs.160483 /len=3035	40419_at
PRG2 (proteoglycan 2, bone marrow (natural killer cell activator, eosinophil granule major basic	Z26248	Hs.99962	NM_002728	11q12	Cluster Incl. Z26248:H.sapiens mRNA for eosinophil granule major basic protein /cds=(857,1525) /gb=Z26248 /gi=940510 /ug=Hs.99962 /len=1637	39179_at
IL18RAP (interleukin 18 receptor accessory protein)	AF077346	Hs.158315	NM_003853	2p24.3-p24.1	Cluster Incl. AF077346:Homo sapiens interleukin-18 receptor accessory protein-like mRNA, complete cds /cds=(483,2282) /gb=AF077346 /gi=3851059 /ug=Hs.158315 /len=2681	33093_at

FBN1 (fibrillin 1 (Marfan syndrome))	X63556	Hs.750	NM_000138	15q21.1	Cluster Incl. X63556:H.sapiens mRNA for fibrillin /cds=(0,9010) /gb=X63556 /gi=397553 /ug=Hs.750 /len=9927	32535_at
LSP1 (lymphocyte-specific protein 1)	M33552	Hs.56729	NM_002339	11p15.5	Cluster Incl. M33552:Human lymphocyte-specific protein 1 (LSP1) mRNA, complete cds /cds=(108,1127) /gb=M33552 /gi=187237 /ug=Hs.56729 /len=1631	36493_at
VNN2 (vanin 2)	D89974	Hs.121102	NM_004665	6q23-q24	Cluster Incl. D89974:Homo sapiens mRNA for glycosylphosphatidyl inositol-anchored protein GPI-80, complete cds /cds=(11,1573) /gb=D89974 /gi=5541649 /ug=Hs.121102 /len=2004	34498_at
PIR121(p53 inducible protein)	L47738	Hs.258503		5	Cluster Incl. L47738:Homo sapiens inducible protein mRNA, complete cds /cds=(1004,1714) /gb=L47738 /gi=1009098 /ug=Hs.80313 /len=2881	37579_at

BASP1 (brain abundant, membrane attached signal protein 1)	AF039656	Hs.79516	NM_006317	5p15.1-p14	Cluster Incl. AF039656:Homo sapiens neuronal tissue-enriched acidic protein (NAP-22) mRNA, complete cds /cds=(52,735) /gb=AF039656 /gi=2773159 /ug=Hs.79516 /len=1467	32607_at
CSF1 (colony stimulating factor 1 (macrophage))	M37435	Hs.173894	NM_000757	1p21-p13	M37435 /FEATURE= /DEFINITION=HUMCSDF1 Human macrophage-specific colony-stimulating factor (CSF-1) mRNA, complete cds	882_at
KIAA0370(KIAA0370 protein)	AB002368			16	Cluster Incl. AB002368:Human mRNA for KIAA0370 gene, partial cds /cds=(0,2406) /gb=AB002368 /gi=2224680 /ug=Hs.70500 /len=5724	35630_at
LILRA3 (leukocyte immunoglobulin-like receptor, subfamily A (without TM domain), member 3)	AF025527	Hs.113277	NM_006865	19q13.4	Cluster Incl. AF025527:Homo sapiens leucocyte immunoglobulin-like receptor-4 (LIR-4) mRNA, complete cds /cds=(93,1412) /gb=AF025527 /gi=2653860 /ug=Hs.113277 /len=1606	35094_f_at

CHRNE (cholinergic receptor, nicotinic, epsilon polypeptide)	X66403	Hs.278295	NM_000080	17p13-p12	Cluster Incl. X66403:H.sapiens mRNA for acetylcholine receptor (epsilon subunit) /cds=(11,1492) /gb=X66403 /gi=560152 /ug=Hs.112028 /len=2457	39834_at
[HSPC022(HSPC022 protein)	W68830	Hs.301175	NM_014029	22	Cluster Incl. W68830:zd37g06.r1 Homo sapiens cDNA, 5 end /clone=IMAGE-342874 /clone_end=5 - /gb=W68830 /gi=1377739 /ug=Hs.173466 /len=614	32736_at
AHCP(Autosomal Highly Conserved Protein)	AL050128	Hs.95260	NM_016255	6	Cluster Incl. AL050128:Homo sapiens mRNA; cDNA DKFZp586G051 (from clone DKFZp586G051) /cds=UNKNOWN /gb=AL050128 /gi=4884335 /ug=Hs.95260 /len=1950	38318_at
MACS (myristoylated alanine-rich protein kinase C substrate (MARCKS, 80K-L))	D10522	Hs.75607	NM_002356	6q22.2	Cluster Incl. D10522:Homo sapiens mRNA for 80K-L protein, complete cds /cds=(369,1367) /gb=D10522 /gi=219893 /ug=Hs.75607 /len=2589	32434_at

CBX7 (chromobox homolog 7)	AL031846			22q13.1	Cluster Incl. AL031846:dJ742C19.5 (novel Chromobox protein) /cds=(89,844) /gb=AL031846 /gi=4164368 /ug=Hs.7442 /len=3964	36894_at
MGAM (maltase-glucoamylase (alpha-glucosidase))	AF016833	Hs.122785	NM_004668	7	Cluster Incl. AF016833:Homo sapiens maltase-glucoamylase mRNA, complete cds /cds=(54,5627) /gb=AF016833 /gi=2826520 /ug=Hs.122785 /len=6483	34509_at
GCA (grancalcin, EF-hand calcium-binding protein)	M81637	Hs.79381	NM_012198	2p14-q14.3	Cluster Incl. M81637:Human grancalcin mRNA, complete cds /cds=(119,772) /gb=M81637 /gi=183030 /ug=Hs.79381 /len=1652	37556_at
TALDO1 (transaldolase 1)	AF010400	Hs.77290	NM_006755	11p15.5-p15.4	Cluster Incl. AF010400:untitled /cds=(50,1063) /gb=AF010400 /gi=2612878 /ug=Hs.77290 /len=1242	37311_at
CPT1B (carnitine palmitoyltransferase I, muscle)	Y08683	Hs.29331	NM_004377	22q13.33	Cluster Incl. Y08683:H.sapiens mRNA for carnitine palmitoyltransferase I type II /cds=(51,2369) /gb=Y08683 /gi=1871536	35935_at

					/ug=Hs.211565 /len=2624	
KIAA1109(KIAA1109 protein)	AB029032	Hs.6606		4	Cluster Incl. AB029032:Homo sapiens mRNA for KIAA1109 protein, partial cds /cds=(0,5873) /gb=AB029032 /gi=5689554 /ug=Hs.6606 /len=6377	36814_at
PSTPIP1 (proline-serine-threonine phosphatase interacting protein 1)	U94778	Hs.129758	NM_003978	15q24-q25.1	Cluster Incl. U94778:Human PEST phosphatase interacting protein homolog (H-PIP) mRNA, complete cds /cds=(216,1466) /gb=U94778 /gi=4100161 /ug=Hs.129758 /len=1656	34914_at
PPBP (pro-platelet basic protein (includes platelet basic protein, beta-thromboglobulin, connective	M54995	Hs.2164	NM_002704	4q12-q13	Cluster Incl. M54995:Human connective tissue activation peptide III mRNA, complete cds /cds=(66,452) /gb=M54995 /gi=181175 /ug=Hs.2164 /len=673	39208_i_at
ANXA11 (annexin A11)	L19605	Hs.75510	NM_001157	10q22-q23	Cluster Incl. L19605:Homo sapiens 56K autoantigen annexin XI gene mRNA, complete cds /cds=(178,1695) /gb=L19605	36637_at

					/gi=457128 /ug=Hs.75510 /len=1958	
PGLYRP (peptidoglycan recognition protein)	AF076483	Hs.137583	NM_005091	19q13.2-q13.3	Cluster Incl. AF076483:Homo sapiens peptidoglycan recognition protein precursor (PGRP) mRNA, complete cds /cds=(44,634) /gb=AF076483 /gi=3342532 /ug=Hs.137583 /len=690	31381_at
SDF2 (stromal cell-derived factor 2)	D50645	Hs.118684	NM_006923	17q11.2	Cluster Incl. D50645:Homo sapiens mRNA for SDF2, complete cds /cds=(39,674) /gb=D50645 /gi=1741867 /ug=Hs.118684 /len=1085	41627_at
PPBP (pro-platelet basic protein (includes platelet basic protein, beta-thromboglobulin, connective	M54995	Hs.2164	NM_002704	4q12-q13	Cluster Incl. M54995:Human connective tissue activation peptide III mRNA, complete cds /cds=(66,452) /gb=M54995 /gi=181175 /ug=Hs.2164 /len=673	39209_r_at
tissue-activating peptide III, neutrophil-activating peptide-2))					Cluster Incl. AB002369:Human mRNA for KIAA0371 gene, complete cds /cds=(247,3843) /gb=AB002369	35739_at

					/gi=2224682 /ug=Hs.63302 /len=5886	
TNFRSF12 (tumor necrosis factor receptor superfamily, member 12 (translocating chain-association membrane protein))	U83600	Hs.180338	NM_003790	1p36.2	U83600 /FEATURE=mRNA /DEFINITION=HSU83600 Human death domain receptor 3 (DDR3) mRNA, alternatively spliced form 2, partial cds	1210_s_at
KCNQ2 (potassium voltage-gated channel, KQT-like subfamily, member 2)	Y15065	Hs.4975	NM_004518	20q13.3	Cluster Incl. Y15065:Homo sapiens mRNA for voltage gated potassium channel /cds=(42,2576) /gb=Y15065 /gi=2826772 /ug=Hs.4975 /len=7407	41589_at
INHA (inhibin, alpha)	M13981	Hs.1734	NM_002191	2q33-q36	M13981 /FEATURE= /DEFINITION=HUMINHA Human inhibin A-subunit mRNA, complete cds	255_s_at
RAC1 (ras-related C3 botulinum toxin substrate 1 (rho family, small GTP binding protein Rac1))	D25274	Hs.173737	NM_006908	7p22	Cluster Incl. D25274:Homo sapiens mRNA, clone-PO2ST9 /cds=UNKNOWN /gb=D25274 /gi=464185 /ug=Hs.173737 /len=1232	40864_at

KIAA0014(KIAA0014 gene product)	D25216	Hs.155650	NM_014665	8	Cluster Incl. D25216:Human mRNA for KIAA0014 gene, complete cds /cds=(146,1627) /gb=D25216 /gi=434774 /ug=Hs.155650 /len=5323	32062_at
MLH1 (mutL (E. coli) homolog 1 (colon cancer, nonpolyposis type 2))	AF001359	Hs.57301	NM_000249	3p21.3	AF001359 /FEATURE= /DEFINITION=AF001359 Homo sapiens DNA mismatch repair protein (hMLH1) mRNA, alternatively spliced, partial cds	1944_f_at
NMP200(nuclear matrix protein NMP200 *related to splicing factor* PRP19)	AI761148	Hs.173980	NM_014502	11	Cluster Incl. AI761148:wh97h07.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-2388733 /clone_end=3 /gb=AI761148 /gi=5176815 /ug=Hs.173980 /len=443	33231_at
CST3 (cystatin C (amyloid angiopathy and cerebral hemorrhage))	AI362017	Hs.135084	NM_000099	20p11.2	Cluster Incl. AI362017:qy39a10.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-2014362 /clone_end=3 /gb=AI362017 /gi=4113638 /ug=Hs.135084 /len=778	39689_at

EIF3S2 (eukaryotic translation initiation factor 3, subunit 2 (beta, 36kD))	U39067	Hs.192023	NM_003757	1p34.1	Cluster Incl. U39067:Homo sapiens translation initiation factor eIF3 p36 subunit mRNA, complete cds /cds=(17,994) /gb=U39067 /gi=1718194 /ug=Hs.192023 /len=1402	32230_at
CNIL(cornichon homolog (Drosophila))	AF104398	Hs.201673	NM_005776	14	Cluster Incl. AF104398:Homo sapiens cornichon mRNA, complete cds /cds=(56,490) /gb=AF104398 /gi=4063708 /ug=Hs.201673 /len=1379	32803_at
SERPINA6 (serine (or cysteine) proteinase inhibitor, clade A (alpha-1 antiproteinase,antitrypsin), member 6)	J02943	Hs.1305	NM_001756	14q32.1	Cluster Incl. J02943:Human corticosteroid binding globulin mRNA, complete cds /cds=(35,1252) /gb=J02943 /gi=179970 /ug=Hs.1305 /len=1422	37833_at
RAC1 (ras-related C3 botulinum toxin substrate 1 (rho family, small GTP binding protein Rac1))	M29870	Hs.173737	NM_006908	7p22	M29870 /FEATURE= /DEFINITION=HUMRACA Human ras-related C3 botulinum toxin substrate (rac) mRNA, complete cds	2050_s_at

PLOD3 (procollagen-lysine, 2-oxoglutarate 5-dioxygenase 3)	AF046889	Hs.153357	NM_001084	7q22	Cluster Incl. AF046889:Homo sapiens lysyl hydroxylase isoform 3 (PLOD3) mRNA, complete cds /cds=(216,2432) /gb=AF046889 /gi=3153234 /ug=Hs.153357 /len=2735	39801_at
SET (SET translocation (myeloid leukemia-associated))	M93651	Hs.145279	NM_003011	9q34	Cluster Incl. M93651:Human set gene, complete cds /cds=(3,836) /gb=M93651 /gi=338038 /ug=Hs.145279 /len=2562	40189_at
GPX1 (glutathione peroxidase 1)	X13710	Hs.76686	NM_000581	3p21.3	Cluster Incl. X13710:H.sapiens unspliced mRNA for glutathione peroxidase /cds=UNKNOWN /gb=X13710 /gi=35387 /ug=Hs.76686 /len=1100	37033_s_at
SAM68(src associated in mitosis, 68 kDa)	M88108	Hs.119537	NM_006559	1	Cluster Incl. M88108:Human p62 mRNA, complete cds /cds=(106,1437) /gb=M88108 /gi=189499 /ug=Hs.119537 /len=2685	39346_at
CD5 (CD5 antigen (p56-62))	X04391	Hs.58685	NM_014207	11q13	Cluster Incl. X04391:Human mRNA for lymphocyte glycoprotein T1/Leu-1	32953_at

					/cds=(72,1559) /gb=X04391 /gi=37186 /ug=Hs.234745 /len=2320	
APLP2 (amyloid beta (A4) precursor-like protein 2)	S60099	Hs.279518	NM_001642	11q24	Cluster Incl. S60099:APPH=amyloid precursor protein homolog [human, placenta, mRNA, 3727 nt] /cds=(72,2363) /gb=S60099 /gi=300168 /ug=Hs.64797 /len=3727	33944_at
ZNF212 (zinc finger protein 212)	U38864	Hs.108139	NM_012256	7q36.1	Cluster Incl. U38864:Human zinc-finger protein C2H2-150 mRNA, complete cds /cds=(220,1065) /gb=U38864 /gi=1055340 /ug=Hs.108139 /len=2235	41426_at
MTVR(Mouse Mammary Tumor Virus Receptor homolog)	AF052151	Hs.18686			Cluster Incl. AF052151:Homo sapiens clone 24574 mRNA sequence /cds=UNKNOWN /gb=AF052151 /gi=3360461 /ug=Hs.18686 /len=1337	32209_at
LOC56928(hypothetical protein from EUROIMAGE 42353)	AC004410	Hs.284161		19	Cluster Incl. AC004410:Homo sapiens *chromosome* 19, *fosmid* 39554 /cds=(0,1196) /gb=AC004410 /gi=2959558	35426_at

					/ug=Hs.167352 /len=1197	
					Cluster Incl. W28227:43h1 Homo sapiens cDNA /gb=W28227 /gi=1308175 /ug=Hs.167985 /len=843	40558_at
DNPEP (aspartyl aminopeptidase)	AF005050	Hs.258551	NM_012100	2q35	Cluster Incl. AF005050:Homo sapiens aspartyl aminopeptidase mRNA, complete cds /cds=(170,1588) /gb=AF005050 /gi=4101588 /ug=Hs.108117 /len=1694	38703_at
KIAA0911(calsyntenin 1)	AB020718	Hs.29665	NM_014944	1	Cluster Incl. AB020718:Homo sapiens mRNA for KIAA0911 protein, complete cds /cds=(793,3738) /gb=AB020718 /gi=4240310 /ug=Hs.29665 /len=5219	41498_at
UQCRC2 (ubiquinol-cytochrome c reductase core protein II)	J04973	Hs.173554	NM_003366	16p12	Cluster Incl. J04973:Human cytochrome bc-1 complex core protein II mRNA, complete cds /cds=(53,1414) /gb=J04973 /gi=180927 /ug=Hs.173554 /len=1588	40854_at

P2RX4 (purinergic receptor P2X, ligand-gated ion channel, 4)	U83993	Hs.321709	NM_002560	12q24.32	Cluster Incl. U83993:Human P2X4 purinoreceptor mRNA, complete cds /cds=(309,1475) /gb=U83993 /gi=4099120 /ug=Hs.9610 /len=2031	38332_at
CLCN7 (chloride channel 7)	Z67743	Hs.80768	NM_001287	16p13	Cluster Incl. Z67743:H.sapiens mRNA for CLC-7 chloride channel protein /cds=(0,2369) /gb=Z67743 /gi=1177439 /ug=Hs.80768 /len=2393	38069_at
NME2 (non-metastatic cells 2, protein (NM23B) expressed in)	X58965	Hs.275163	NM_002512	17q21.3	X58965 /FEATURE= /DEFINITION=HSNM23H2G H.sapiens RNA for nm23-H2 gene	1980_s_at

Table 9:

UCL/HGNC/HUGO Human Gene Nomenclature Database Symbol	GenBank Accession No.	UniGene Cluster	RefSeq	Chromosomal Location	Description Unigene Build #95	Gene Name
					c /FEATURE= /DEFINITION=HUM4AI Human mRNA for eukaryotic initiation factor 4AI	1199_at
MPB1 (MYC promoter-binding protein 1)	M55914	Hs.284127	NM_005945	1pter-p35	M55914 /FEATURE= /DEFINITION=HUMCMYCQ Human c-myc binding protein (MBP-1) mRNA, complete cds	2035_s_at
TXN (thioredoxin)	AI653621	Hs.76136	NM_003329	9q31	Cluster Incl. AI653621:tz21b11.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-2289213 /clone_end=3 /gb=AI653621 /gi=4737600 /ug=Hs.76136 /len=598	36992_at

RNASE2 (ribonuclease, RNase A family, 2 (liver, eosinophil-derived neurotoxin))	X55988	Hs.728	NM_002934	14q24-q31	Cluster Incl. X55988:Human EDN mRNA for eosinophil derived neurotoxin /cds=(71,556) /gb=X55988 /gi=31088 /ug=Hs.728 /len=735	36766_at
RNASE2 (ribonuclease, RNase A family, 2 (liver, eosinophil-derived neurotoxin))	X95735	Hs.728	NM_002934	14q24-q31	Cluster Incl. X95735:Homo sapiens mRNA for zyxin /cds=(71,1789) /gb=X95735 /gi=1545953 /ug=Hs.75873 /len=2217	36958_at
NME2 (non-metastatic cells 2, protein (NM23B) expressed in)	X58965	Hs.275163	Hs.275163	17q21.3	X58965 /FEATURE= /DEFINITION=HSNM23H2G H.sapiens RNA for nm23-H2 gene	1980_s_at
H2AFY (H2A histone family, member Y)	AF054174	Hs.75258	NM_004893	5q31.3-q32	Cluster Incl. AF054174:Homo sapiens histone macroH2A1.2 mRNA, complete cds /cds=(173,1288) /gb=AF054174 /gi=3341991 /ug=Hs.75258 /len=1881	36576_at
IDH2 (isocitrate dehydrogenase 2 (NADP+), mitochondrial)	X69433	Hs.5337	Hs.5337	15q26.1	Cluster Incl. X69433:H.sapiens mRNA for mitochondrial isocitrate dehydrogenase (NADP+) /cds=(86,1444) /gb=X69433	32332_at

					/gi=872120 /ug=Hs.182740 /len=1751	
TST (thiosulfate sulfurtransferase (rhodanese))	X59434	Hs.248267	NM_003312	22q13.1	Cluster Incl. X59434:Human rohu mRNA for rhodanese /cds=(34,924) /gb=X59434 /gi=432375 /ug=Hs.74097 /len=1232	36124_at
PGD (phosphogluconate dehydrogenase)	U30255	Hs.75888	NM_002631	1p36.3-p36.13	Cluster Incl. U30255:Human phosphogluconate dehydrogenase (hPGDH) gene, complete cds /cds=(6,1457) /gb=U30255 /gi=984324 /ug=Hs.75888 /len=1536	36963_at
PSMA6 (proteasome (prosome, macropain) subunit, alpha type, 6)	X59417	Hs.336907		14q13	Cluster Incl. X59417:H.sapiens PROS-27 mRNA /cds=(62,802) /gb=X59417 /gi=35681 /ug=Hs.74077 /len=964	36122_at
CD63 (CD63 antigen (melanoma 1 antigen)	X62654	Hs.76294	NM_001780	12q12-q13	Cluster Incl. X62654:H.sapiens gene for Me491/CD63 antigen /cds=(69,785) /gb=X62654 /gi=430755 /ug=Hs.76294 /len=873	37003_at

ALDOA (aldolase A, fructose-bisphosphate)	X05236	Hs.273415	NM_000034	16q22-q24	Cluster Incl. X05236:Human fibroblast mRNA for aldolase A /cds=(146,1240) /gb=X05236 /gi=28596 /ug=Hs.183760 /len=1440	32336_at
CREG (cellular repressor of E1A-stimulated genes)	AF084523	Hs.5710	Hs.5710	1q24	Cluster Incl. AF084523:Homo sapiens cellular repressor of E1A-stimulated genes CREG mRNA, complete cds /cds=(33,695) /gb=AF084523 /gi=3550342 /ug=Hs.5710 /len=1974	35311_at
NME4 (non-metastatic cells 4, protein expressed in)	Y07604	Hs.9235	Hs.9235	16p13.3	Cluster Incl. Y07604:H.sapiens mRNA for nucleoside-diphosphate kinase /cds=(11,574) /gb=Y07604 /gi=1945761 /ug=Hs.9235 /len=879	39089_at
CUTL1 (cut (Drosophila)-like 1 (CCAAT displacement protein))	L12579	Hs.147049	NM_001913	7q22	Cluster Incl. L12579:Human alternatively spliced CUTL1 mRNA, complete cds /cds=(19,2055) /gb=L12579 /gi=457516 /ug=Hs.147049 /len=2855	31822_at

LDHA (lactate dehydrogenase A)	X02152	Hs.2795	NM_005566	11p15.4	Cluster Incl. X02152:Human mRNA for lactate dehydrogenase-A (LDH-A, EC 1.1.1.27) /cds=(97,1095) /gb=X02152 /gi=34312 /ug=Hs.2795 /len=1661	41485_at
PTTG1IP (pituitary tumor-transforming 1 interacting protein)	Z50022	Hs.111126	NM_004339	21q22.3	Cluster Incl. Z50022:H.sapiens mRNA for surface glycoprotein /cds=(93,635) /gb=Z50022 /gi=1107702 /ug=Hs.111126 /len=2617	39003_at
H2AV(histone H2A.F/Z variant)	AW007731	Hs.301005	NM_012412	7	Cluster Incl. AW007731:wt68d11.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-2512629 /clone_end=3 /gb=AW007731 /gi=5856509 /ug=Hs.9242 /len=659	39092_at
RAB32 (RAB32, member RAS oncogene family)	U59878	Hs.32217	NM_006834	6	Cluster Incl. U59878:Human low-Mr GTP-binding protein (RAB32) mRNA, partial cds /cds=(0,632) /gb=U59878 /gi=1388196 /ug=Hs.32217 /len=980	41523_at

DDAH2 (dimethylarginine dimethylaminohydrolase 2)	AJ012008	Hs.247362	NM_013974	6p21.3	Cluster Incl. AJ012008:Homo sapiens genes *encoding* RNCC protein, DDAH protein, Ly6-C protein, Ly6-D protein and immunoglobulin receptor /cds=(218,943) /gb=AJ012008 /gi=5304874 /ug=Hs.74276 /len=1200	38131_at
GNG5 (guanine nucleotide binding protein (G protein), gamma 5)	AI541042	Hs.5322	NM_005274	1p22	Cluster Incl. AI541042:pec1.2-1.D12.r Homo sapiens cDNA, 5 end /clone_end=5 /gb=AI541042 /gi=4458415 /ug=Hs.5322 /len=688	35272_at
GAPD (glyceraldehyde-3-phosphate dehydrogenase)	M33197	Hs.169476	Hs.169476	12p13	Homo sapiens /REF=M33197 /DEF=Human glyceraldehyde-3-phosphate dehydrogenase (GAPDH) mRNA, *complete* cds /LEN=1268 (_5, _M, _3 represent transcript regions 5 prime, *Middle*, and 3 prime respectively)	AFFX-HUMGAPDH/M33197_5_at
GAPD (glyceraldehyde-3-phosphate dehydrogenase)	M33197	Hs.169476	NM_002046	12p13	Homo sapiens /REF=M33197 /DEF=Human glyceraldehyde-3-phosphate dehydrogenase (GAPDH) mRNA, complete cds /LEN=1268 (_5, _M, _3	AFFX-HUMGAPDH/M33197_M_at

					complete cds /LEN=1268 (_5, _M, _3 represent transcript regions 5 prime, Middle, and 3 prime respectively)	M_at
GAPD (glyceraldehyde-3-phosphate dehydrogenase)	M33197	Hs.169476	NM_002046	12p13	Homo sapiens /REF=M33197 /DEF=Human glyceraldehyde-3-phosphate dehydrogenase (GAPDH) mRNA, complete cds /LEN=1268 (_5, _M, _3 represent transcript regions 5 prime, Middle, and 3 prime respectively)	AFFX-HUMGAPDH/M33197_3_at
CST3 (cystatin C (amyloid angiopathy and cerebral hemorrhage))	AI362017	Hs.135084	NM_000099	20p11.2	Cluster Incl. AI362017:qy39a10.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-2014362 /clone_end=3 /gb=AI362017 /gi=4113638 /ug=Hs.135084 /len=778	39689_at
USP5 (ubiquitin specific protease 5 (isopeptidase T))	U47924	Hs.3759	NM_003481	12p13	Cluster Incl. U47924:Human chromosome 12p13 sequence /cds=(373,1122) /gb=U47924 /gi=1633547 /ug=Hs.83848 /len=1843	34003_at

MAPKAPK3 (mitogen-activated protein kinase-activated protein kinase 3)	U09578	Hs.227789	NM_004635	3p21.3	U09578 /FEATURE= /DEFINITION=HSU09578 Homo sapiens MAPKAP kinase (3pK) mRNA, complete cds	1637_at
ARHG (ras homolog gene family, member G (rho G))	X61587	Hs.75082	NM_001665	11p15.5-p15.4	Cluster Incl. X61587:H.sapiens rhoG mRNA for GTPase /cds=(129,704) /gb=X61587 /gi=36035- /ug=Hs.75082 /len=1284	36902_at
TALDO1 (transaldolase 1)	AF010400	Hs.77290	NM_006755	11p15.5-p15.4	Cluster Incl. AF010400:untitled /cds=(50,1063) /gb=AF010400 /gi=2612878 /ug=Hs.77290 /len=1242	37311_at
HNRPAB (heterogeneous nuclear ribonucleoprotein A/B)	M65028	Hs.81361	NM_004499	5q35	Cluster Incl. M65028:Human hnRNP type A/B protein mRNA, complete cds /cds=(142,996) /gb=M65028 /gi=337450 /ug=Hs.81361 /len=1537	38094_at
FAH (fumarylacetoacetate hydrolase (fumarylacetoacetase))	M55150	Hs.73875	NM_000137	15q23-q25	Cluster Incl. M55150:Human fumarylacetoacetate hydrolase mRNA, complete cds /cds=(56,1315) /gb=M55150	36876_at

					/gi=182392 /ug=Hs.73875 /len=1447	
PRG1 (proteoglycan 1, secretory granule)	X17042	Hs.278687	NM_002727	19q13.2	Cluster Incl. X17042:Human mRNA for hematopoetic proteoglycan core protein /cds=(24,500) /gb=X17042 /gi=32432 /ug=Hs.1908 /len=1182	32227_at
M11S1 (membrane component, chromosome 11, surface marker 1)	Z48042	Hs.278672	NM_005898	11p13	Cluster Incl. Z48042:H.sapiens mRNA encoding GPI-anchored protein p137 /cds=(201,2150) /gb=Z48042 /gi=662993 /ug=Hs.101025 /len=3268	39471_at
ATP6F (ATPase, H+ transporting, lysosomal (vacuolar proton pump) 21kD)	D89052	Hs.7476	NM_004047	1p32.3	Cluster Incl. D89052:Homo sapiens mRNA for proton-ATPase-like protein, complete cds /cds=(82,699) /gb=D89052 /gi=1694672 /ug=Hs.7476 /len=987	36167_at
ADAM15 (a disintegrin and metalloproteinase domain 15 (metargidin))	U41767	Hs.92208	NM_003815	1q21.3	Cluster Incl. U41767:Human metargidin precursor mRNA, complete cds /cds=(7,2451) /gb=U41767 /gi=1235673 /ug=Hs.92208 /len=2725	38282_at

COX6A1 (cytochrome c oxidase subunit VIa polypeptide 1)	AI540925	Hs.180714	NM_004373	12q24.2	Cluster Incl. AI540925:PEC1.2_15_A02.r Homo sapiens cDNA, 5 end /clone_end=5 /gb=AI540925 /gi=4458298 /ug=Hs.180714 /len=777	41206_r_at
NFIL3 (*nuclear factor*, interleukin 3 regulated)	X64318	Hs.79334	NM_005384	9q22	Cluster Incl. X64318:H.sapiens E4BP4 gene /cds=(213,1601) /gb=X64318 /gi=30955 /ug=Hs.79334 /len=1904	37544_at
COX8 (cytochrome c oxidase subunit VIII)	AI525665	Hs.81097	NM_004074	11q12-q13	Cluster Incl. AI525665:PT1.3_04_D06.r Homo sapiens cDNA, 5 end /clone_end=5 /gb=AI525665 /gi=4439800 /ug=Hs.81097 /len=834	38080_at
COMT (catechol-O-methyltransferase)	M58525	Hs.240013	NM_000754	22q11.21	Cluster Incl. :Homo sapiens catechol-O-methyltransferase (COMT) mRNA, complete cds /cds=(204,1019) /gb=M58525 /gi=179954 /ug=Hs.78534 /len=1206	34651_at

SCGF (stem cell growth factor; lymphocyte secreted C-type lectin)	AF020044	Hs.105927	NM_002975	19q13.3	Cluster Incl. AF020044:Homo sapiens lymphocyte secreted C-type lectin precursor, mRNA, complete cds /cds=(179,1150) /gb=AF020044 /gi=2828595 /ug=Hs.105927 /len=1391	37147_at
P100(staphylococcal nuclease domain containing 1)]	U22055	19	NM_014390	7	Cluster Incl. U22055:Human 100 kDa coactivator mRNA, complete cds /cds=(267,2924) /gb=U22055 /gi=799176 /ug=Hs.79093 /len=3480	37730_at
GPSN2 (glycoprotein, synaptic 2)	AF038958	Hs.306122	NM_004868	19p13.2	Cluster Incl. AF038958:Homo sapiens *synaptic glycoprotein* SC2 spliced variant mRNA, complete cds /cds=(76,1002) /gb=AF038958 /gi=3329385 /ug=Hs.109051 /len=1116	38966_at
DDOST (dolichyl-diphosphooligosaccharide-protein glycosyltransferase)	D29643	Hs.34789	NM_005216	1p36.1	Cluster Incl. D29643:Human mRNA for KIAA0115 gene, complete cds /cds=(106,1476) /gb=D29643 /gi=473936 /ug=Hs.89674 /len=1668	38791_at

HDLBP (high density lipoprotein binding protein (vigilin))	M64098	Hs.177516	NM_005336	2q37	Cluster Incl. M64098:Human high density lipoprotein binding protein (HBP) mRNA, complete cds /cds=(154,3960) /gb=M64098 /gi=183891 /ug=Hs.177516 /len=4354	31504_at
LOC57019(hypothetical protein)	AC004382	Hs.4900	NM_020313	16	Cluster Incl. AC004382:Homo sapiens Chromosome 16 BAC clone CIT987SK-A-152E5 /cds=(0,935) /gb=AC004382 /gi=3252819 /ug=Hs.79402 /len=1659	32600_at
RAB13 (RAB13, member RAS oncogene family)	X75593	Hs.151536	NM_002870	12q13	Cluster Incl. X75593:H.sapiens mRNA for rab 13 /cds=(139,750) /gb=X75593 /gi=452319 /ug=Hs.151536 /len=1238	40210_at
PFN1 (profilin 1)	J03191	Hs.75721	NM_005022	17p13.3	Cluster Incl. J03191:Human profilin mRNA, complete cds /cds=(127,549) /gb=J03191 /gi=190385 /ug=Hs.75721 /len=793	36675_r_at
DXS1357E(accessory proteins BAP31/BAP29)	X81817			X	Cluster Incl. X81817:H.sapiens BAP31 mRNA /cds=(73,813) /gb=X81817	41724_at

					/gi=550342 /ug=Hs.181373 /len=1504	
RAC1 (ras-related C3 botulinum toxin substrate 1 (rho family, small GTP binding protein Rac1)	M29870	Hs.173737	NM_006908	7p22	M29870 /FEATURE= /DEFINITION=HUMRACA Human ras-related C3 botulinum toxin substrate (rac) mRNA, complete cds	2050_s_at
FKBP1A (FK506-binding protein 1A (12kD))	M34539	Hs.179661	NM_000801	20p13	M34539 /FEATURE= /DEFINITION=HUMFKBP Human FK506-binding protein (FKBP) mRNA, complete cds	880_at
YF13H12(protein expressed in thyroid)	D83198	Hs.7486	NM_014297	19	Cluster Incl. D83198:Homo sapiens mRNA expressed in thyroid gland /cds=(341,1024) /gb=D83198 /gi=3893156 /ug=Hs.7486 /len=1154	36170_at
ANXA1 (annexin A1)	X05908	Hs.78225	NM_000700	9q12-q21.2	Cluster Incl. X05908:Human mRNA for lipocortin /cds=(74,1114) /gb=X05908 /gi=34387 /ug=Hs.78225 /len=1399	37403_at

Approved UCL/HGNC/HUGO Human Gene Nomenclature database symbol	AF046889	Hs.153357	NM_001084	7q22	Cluster Incl. AF046889:Homo sapiens lysyl hydroxylase isoform 3 (PLOD3) mRNA, complete cds /cds=(216,2432) /gb=AF046889 /gi=3153234 /ug=Hs.153357 /len=2735	39801_at
SPN (sialophorin (gpL115, leukosialin, CD43))	J04168	Hs.80738	NM_003123	16p11.2	Cluster Incl. J04168:Human leukosialin mRNA, complete cds . /cds=(95,1297) /gb=J04168 /gi=187118 /ug=Hs.80738 /len=2288	36798_g_at
RAGD(Rag D protein)]	W27549	Hs.238679	NM_021244	6	Cluster Incl. W27549:32d11 Homo sapiens cDNA /gb=W27549 /gi=1307353 /ug=Hs.235634 /len=912	32963_s_at
IMPDH2 (IMP (inosine monophosphate) dehydrogenase 2)	L33842	Hs.75432	NM_000884	3p21.2	Cluster Incl. L33842:Homo sapiens (clone FFE-7) type II inosine monophosphate dehydrogenase (IMPDH2) gene, exons 1-13, complete cds /cds=(102,1646) /gb=L33842 /gi=602457 /ug=Hs.75432 /len=1688	36624_at

PSMD9 (proteasome (prosome, macropain) 26S subunit, non-ATPase, 9	AB003177	Hs.5648	NM_002813	12q24.31-q24.32	AB003177 /FEATURE= /DEFINITION=AB003177 Homo sapiens mRNA for proteasome subunit p27, complete cds	1444_at
NSEP1 (nuclease sensitive element binding protein 1)	M85234	Hs.74497	NM_004559	1p34	Cluster Incl. M85234:Human nuclease sensitive element binding protein-1 mRNA, complete cds /cds=(234,1202) /gb=M85234 /gi=337427 /ug=Hs.184712 /len=1474	32340_s_at
LOC95295(hypothetical gene supported by V00599; BC001938; BC007605; BC008791)	V00599			6	V00599 /FEATURE=mRNA /DEFINITION=HSTUB2 Human mRNA fragment encoding beta-tubulin. (from clone D-beta-1)	151_s_at
MAZ (MYC-associated zinc finger protein (purine-binding transcription factor))	M94046	Hs.7647	NM_002383	16p11.2	Cluster Incl. M94046:Human zinc finger protein (MAZ) mRNA /cds=UNKNOWN /gb=M94046 /gi=187393 /ug=Hs.7647 /len=2389	32553_at

HDGF (hepatoma-derived growth factor (high-mobility group protein 1-like))	L24521	Hs.89525	NM_004494	xq25	Cluster Incl. L24521:Human transformation-related protein mRNA, 3 end /cds=(0,1108) /gb=L24521 /gi=403459 /ug=Hs.169225 /len=1240	36446_s_at
KDELR1 (KDEL (Lys-Asp-Glu-Leu) endoplasmic reticulum protein retention receptor 1)	X55885	Hs.78040	NM_006801	19q13.3	Cluster Incl. X55885:Human mRNA for a presumptive KDEL receptor /cds=(146,784) /gb=X55885 /gi=34030 /ug=Hs.78040 /len=1086	37386_i_at
ACADVL (acyl-Coenzyme A dehydrogenase, very long chain)	L46590	Hs.82208	NM_000018	17p13-p11	Cluster Incl. L46590:Homo sapiens very long chain acyl-CoA dehydrogenase gene, exons 1-20, complete cds /cds=(88,2055) /gb=L46590 /gi=1008851 /ug=Hs.82208 /len=2224	38376_at
DGKZ (diacylglycerol kinase, zeta (104kD))	U94905	Hs.89981	NM_003646	11p11.2	Cluster Incl. U94905:Human diacylglycerol kinase zeta mRNA, alternatively spliced, complete cds /cds=(125,3478) /gb=U94905 /gi=2183037 /ug=Hs.89981 /len=4094	38003_s_at

SNRPF (small nuclear *ribonucleoprotein* polypeptide F)	AI032612	Hs.105465	NM_003095	12	Cluster Incl. AI032612:ow17e07.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-1647108 /clone_end=3 /gb=AI032612 /gi=3253738 /ug=Hs.105465 /len=582	41403_at
ATP6S1 (*ATPase, H+ transporting*, lysosomal (vacuolar proton pump), subunit 1)	D16469	Hs.6551	NM_001183	xq28	Cluster Incl. D16469:Human mRNA for ORF, Xq terminal portion /cds=(1353,2198) /gb=D16469 /gi=758583 /ug=Hs.6551 /len=2823	35770_at
TRIM28 (tripartite motif-containing 28)	X97548	Hs.228059	NM_005762	5	Cluster Incl. X97548:H.sapiens mRNA for TIF1beta zinc finger protein /cds=(361,2868) /gb=X97548 /gi=1524108 /ug=Hs.228059 /len=3035	33425_at
K-ALPHA-1(tubulin, alpha, ubiquitous)	K00558	Hs.334842	NM_006082	12	Cluster Incl. K00558:human alpha-tubulin mRNA, complete cds /cds=(67,1422) /gb=K00558 /gi=340020 /ug=Hs.169476 /len=1596	32272_at
COX7B (cytochrome c oxidase subunit VIIb)	N50520	Hs.75752	NM_001866	xp21.1-q21.33	Cluster Incl. N50520:yy89b05.s1 Homo sapiens cDNA, 3 end /clone=IMAGE-	36687_at

					280689 /clone_end=3 /gb=N50520 /gi=1191686 /ug=Hs.75752 /len=550	
CHC1 (chromosome condensation 1)	D00591	Hs.84746	NM_001269	1p36.1	D00591 /FEATURE=exons#7-14 /DEFINITION=HUMRCC1 Homo sapiens RCC1 gene, exons 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, complete cds	1196_at
TYMS (thymidylate synthetase)	D00596	Hs.82962	NM_001071	18p11.32	D00596 /FEATURE=cds /DEFINITION=HUMTS1 Homo sapiens gene for thymidylate synthase, exons 1, 2, 3, 4, 5, 6, 7, complete cds	1505_at
GPX1 (glutathione peroxidase 1)	X13710	Hs.76686	NM_000581	3p21.3	Cluster Incl. X13710:H.sapiens unspliced mRNA for glutathione peroxidase /cds=UNKNOWN /gb=X13710 /gi=35387 /ug=Hs.76686 /len=1100	37033_s_at
IGFBP7 (insulin-like growth factor binding protein 7)	L19182	Hs.119206	NM_001553	4q12	L19182 /FEATURE= /DEFINITION=HUMMAC25X Human MAC25 mRNA, complete cds	2062_at

HMG20B (high-mobility group 20B)	AF072836	Hs.32317	NM_006339	19p13.3	Cluster Incl. AF072836:Homo sapiens Sox-like transcriptional factor mRNA, complete cds /cds=(18,1043) /gb=AF072836 /gi=3329481 /ug=Hs.32317 /len=1232	41526_at
R33729_1(hypothetical protein R33729_1)]	AI828168	Hs.10927		19	Cluster Incl. AI828168:wk32h09.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-2414081 /clone_end=3 /gb=AI828168 /gi=5448839 /ug=Hs.10927 /len=759	38969_at
MAN2B1 (mannosidase, alpha, class 2B, member 1)	U60899	Hs.279854	NM_000528	19cen-q13.1	Cluster Incl. U60899:Human lysosomal alpha-mannosidase (manB) gene /cds=(309,3341) /gb=U60899 /gi=2209014 /ug=Hs.234070 /len=3443	34670_at
PSMB4 (proteasome (prosome, macropain) subunit, beta type, 4)	D26600	Hs.89545	NM_002796	1q21	D26600 /FEATURE= /DEFINITION=HUMPSH3 Human mRNA for proteasome subunit HsN3, complete cds	1311_at

CDK2AP1 (CDK2-associated protein 1)	AF006484	Hs.3436	NM_004642	12q24.31	Cluster Incl. AF006484:Homo sapiens putative oral tumor suppressor protein (doc-1) mRNA, complete cds /cds=(522,869) /gb=AF006484 /gi=2738496 /ug=Hs.3436 /len=1608	41535_at
H2AFZ (H2A histone family, member Z)	M37583	Hs.119192	NM_002106	4q24	Cluster Incl. M37583:Human histone (H2A.Z) mRNA, complete cds /cds=(106,492) /gb=M37583 /gi=184059 /ug=Hs.119192 /len=873	39337_at
KIAA0095(KIAA0095 gene product)]	D42085	Hs.155314	NM_014669	16	Cluster Incl. D42085:Human mRNA for KIAA0095 gene, complete cds /cds=(66,2525) /gb=D42085 /gi=577316 /ug=Hs.155314 /len=2681	40271_at
RPN2 (ribophorin II)	AL031659	Hs.75722	NM_002951	20q12-q13.1	Cluster Incl. AL031659:dJ343K2.2.1 (ribophorin II (isoform 1)) /cds=(284,2179) /gb=AL031659 /gi=4468296 /ug=Hs.75722 /len=2488	36676_at

SLC29A1 (solute carrier family 29 (nucleoside transporters), member 1	U81375	Hs.25450	NM_004955	6p21.1-p21.2	Cluster Incl. U81375:Human placental equilibrative nucleoside transporter 1 (hENT1) mRNA, complete cds /cds=(178,1548) /gb=U81375 /gi=1845344 /ug=Hs.25450 /len=2162-	33901_at
OS-9(amplified in osteosarcoma)	U41635	Hs.76228	NM_006812	12	Cluster Incl. U41635:Human OS-9 precurosor mRNA, complete cds /cds=(85,2088) /gb=U41635 /gi=1322233 /ug=Hs.76228 /len=2736	36996_at
TIP47(cargo selection protein (mannose 6 phosphate receptor binding protein))	AF057140	Hs.140452	NM_005817	19	Cluster Incl. AF057140:Homo sapiens cargo selection protein TIP47 (TIP47) mRNA, complete cds /cds=(74,1378) /gb=AF057140 /gi=3095185 /ug=Hs.140452 /len=1974	40169_at
AF053356-CDS2(hypothetical protein AF053356-CDS2)	AF053356	Hs.296336	NM_022574	7	Cluster Incl. AF053356:Homo sapiens chromosome 7q22 sequence /cds=(253,1275) /gb=AF053356 /gi=3135305 /ug=Hs.91299 /len=1638	38831_f_at

CS (citrate synthase)	AF047042	Hs.239760	NM_004077	12p11-qter	Cluster Incl. AF047042:Homo sapiens citrate synthase mRNA, complete cds /cds=(0,1400) /gb=AF047042 /gi=3288814 /ug=Hs.239760 /len=1401	41314_at
ATP5I (ATP synthase, H+ transporting, mitochondrial F0 complex, subunit e)	AA426364	Hs.85539	NM_007100	4p	Cluster Incl. AA426364:zv61b06.s1 Homo sapiens cDNA, 3 end /clone=IMAGE-758099 /clone_end=3 /gb=AA426364 /gi=2106690 /ug=Hs.85539 /len=401	38751_i_at
KIAA0233(KIAA0233 gene product)	D87071	Hs.79077	NM_014745	16	Cluster Incl. D87071:Human mRNA for KIAA0233 gene, complete cds /cds=(2,6109) /gb=D87071 /gi=1510142 /ug=Hs.79077 /len=6368	37281_at
UQCRFS1 (ubiquinol-cytochrome c reductase, Rieske iron-sulfur polypeptide 1)	L32977	Hs.3712	NM_006003	19q12-q13	Cluster Incl. L32977:Homo sapiens (clone f17252) ubiquinol cytochrome c reductase Rieske iron-sulphur protein (UQCRFS1) gene /cds=(90,914) /gb=L32977 /gi=488298 /ug=Hs.3712 /len=1203	34401_at

PSMB7 (proteasome (prosome, macropain) subunit, beta type, 7)	D38048	Hs.118065	NM_002799	9q34.11-q34.12	D38048 /FEATURE= /DEFINITION=D38048 Human mRNA for proteasome subunit z, complete cds	1313_at
YWHAE (tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, epsilon polypeptide)	U54778	Hs.79474	NM_006761	17p13.3	U54778 /FEATURE= /DEFINITION=HSU54778 Human 14-3-3 epsilon mRNA, complete cds	1011_s_at
SMARCA4 (SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4)	U29175	Hs.78202	NM_003072	19p13.2	Cluster Incl. U29175:Human transcriptional activator (BRG1) mRNA, complete cds /cds=(78,5021) /gb=U29175 /gi=902045 /ug=Hs.78202 /len=5247	32579_at
CANX (calnexin)	L10284	Hs.155560	NM_001746	5q35	Cluster Incl. L10284:Homo sapiens integral membrane protein; calnexin, (IP90) mRNA, complete cds /cds=(89,1867) /gb=L10284 /gi=186522 /ug=Hs.155560 /len=4117	40125_at
LIMK2 (LIM domain kinase 2)	AC002073	Hs.278027	NM_005569	22q12.2	Cluster Incl. AC002073:Human PAC clone DJ515N1 from 22q11.2-q22 /cds=(0,2201) /gb=AC002073 /gi=2078469	38618_at

					/ug=Hs.100623 /len=2202	
CAPN4(calpain 4)	X04106	Hs.74451	NM_001749		Cluster Incl. X04106:Human mRNA for calcium dependent protease (small subunit) /cds=(158,964) /gb=X04106 /gi=35327 /ug=Hs.74451 /len=1478	36138_at
DF (D component of complement (adipsin))	M84526	Hs.155597	NM_001928	19	Cluster Incl. M84526:Human adipsin/complement factor D mRNA, complete cds /cds=(54,740) /gb=M84526 /gi=178625 /ug=Hs.155597 /len=1071	40282_s_at
CSF3R (colony stimulating factor 3 receptor (granulocyte))	M59818	Hs.2175	NM_000760	1p35-p34.3	Cluster Incl. M59818:Human granulocyte colony-stimulating factor receptor (G-CSFR-1) mRNA, complete cds /cds=(169,2679) /gb=M59818 /gi=183046 /ug=Hs.2175 /len=2943	34223_at
TNFRSF7 (tumor necrosis factor receptor superfamily, member 7)	M63928	Hs.180841	NM_001242	12p13	Cluster Incl. M63928:Homo sapiens T cell activation antigen (CD27) mRNA, complete cds /cds=(100,882) /gb=M63928	38578_at

					/gi=180084 /ug=Hs.180841 /len=1204	
TRB@ (T cell receptor beta locus)	M12886	Hs.303157		7q35	M12886 /FEATURE= /DEFINITION=HUMTCBYY Human T-cell receptor active beta-chain mRNA, complete cds	1105_s_at
KIAA0275(KIAA0275 gene product)	D87465	Hs.74583	NM_014767	10	Cluster Incl. D87465:Human mRNA for KIAA0275 gene, complete cds /cds=(316,1590) /gb=D87465 /gi=1665814 /ug=Hs.74583 /len=5316	36155_at
TGFBR3 (transforming growth factor, beta receptor III (betaglycan, 300kD))	L07594	Hs.79059	NM_003243	1p33-p32	L07594 /FEATURE= /DEFINITION=HUMTGFB3C Human transforming growth factor-beta type III receptor (TGF-beta) mRNA, complete cds	1897_at
IGHM (immunoglobulin heavy constant mu)	X58529	Hs.302063		14q32.33	Cluster Incl. X58529:Human rearranged immunoglobulin mRNA for mu heavy chain enhancer and constant region /cds=UNKNOWN /gb=X58529 /gi=33480	41166_at

					/ug=Hs.179543 /len=2325	
IGHM (immunoglobulin heavy constant mu)	X67301	Hs.302063		14q32.33	Cluster Incl. X67301:H.sapiens mRNA for IgM heavy chain constant region (Ab63) /cds=(0,1361) /gb=X67301 /gi=38407 /ug=Hs.179543 /len=1453	41164_at
TCL1A (T-cell leukemia/lymphoma 1A)	X82240	Hs.2484	NM_021966	14q32.1	Cluster Incl. X82240:H.sapiens mRNA for Tcell leukemia/lymphoma 1 /cds=(45,389) /gb=X82240 /gi=624960 /ug=Hs.2484 /len=1312	39318_at
PLCE2 (phospholipase C, epsilon 2)	AB029015	Hs.54886		3p25.3-p25.1	Cluster Incl. AB029015:Homo sapiens mRNA for KIAA1092 protein, partial cds /cds=(0,3464) /gb=AB029015 /gi=5689520 /ug=Hs.54886 /len=4147	41796_at
PFTK1 (PFTAIRE protein kinase 1)	AB020641	Hs.57856	NM_012395	7q21-q22	Cluster Incl. AB020641:Homo sapiens mRNA for KIAA0834 protein, complete cds /cds=(144,1499) /gb=AB020641	36502_at

					/gi=4240156 /ug=Hs.57856 /len=4957	
IGHM (immunoglobulin heavy constant mu)	X67301	Hs.302063		14q32.33	Cluster Incl. X67301:H.sapiens mRNA for IgM heavy chain constant region (Ab63) /cds=(0,1361) /gb=X67301 /gi=38407 /ug=Hs.179543 /len=1453	41165_g_at
CBX7 (chromobox homolog 7)	AL031846			22q13.1	Cluster Incl. AL031846:dJ742C19.5 (novel Chromobox protein) /cds=(89,844) /gb=AL031846 /gi=4164368 /ug=Hs.7442 /len=3964	36894_at
DKFZP564K0822(hypothetical protein DKFZp564K0822)	W25986	Hs.4750	NM_030796	7	Cluster Incl. W25986:17e7 Homo sapiens cDNA /gb=W25986 /gi=1306253 /ug=Hs.4750 /len=769	34830_at
BLK (B lymphoid tyrosine kinase)	S76617 /	Hs.2243	NM_001715	8p23-p22	S76617 /FEATURE= /DEFINITION=S76617 blk=protein tyrosine kinase [human, B lymphocytes, mRNA, 2608 nt]	854_at

CD79A (CD79A antigen (immunoglobulin-associated alpha))	U05259	Hs.79630	NM_001783	19q13.2	Cluster Incl. U05259:Human MB-1 gene, complete cds /cds=(36,716) /gb=U05259 /gi=452561 /ug=Hs.79630 /len=1107	38017_at
DGKA (diacylglycerol kinase, alpha (80kD))	X62535	Hs.172690	NM_001345	12q13.3	Cluster Incl. X62535:H.sapiens mRNA for diacylglycerol kinase /cds=(103,2310) /gb=X62535 /gi=30822 /ug=Hs.172690 /len=2564	32716_at
CD19 (CD19 antigen)	M28170	Hs.96023	NM_001770	16p11.2	M28170 /FEATURE= /DEFINITION=HUMCSPC Human cell surface protein CD19 (CD19) gene, complete cds	1096_g_at
SH3BP5 (SH3-domain binding protein 5 (BTK-associated))	AB005047	Hs.109150	NM_004844	1q43	Cluster Incl. AB005047:Homo sapiens mRNA for SH3 binding protein, complete cds /cds=(63,1340) /gb=AB005047 /gi=3116213 /ug=Hs.109150 /len=2570	38968_at
KIAA0226(KIAA0226 gene product)]	D86979	Hs.141296		3	Cluster Incl. D86979:Human mRNA for KIAA0226 gene, complete cds /cds=(622,2877) /gb=D86979 /gi=1504031	31802_at

					/ug=Hs.141296 /len=5891	
NIFU(nitrogen fixation cluster-like)	U47101	Hs.9908		12	Cluster Incl. U47101:Human NifU-like protein (hNifU) mRNA, partial cds /cds=(0,366) /gb=U47101 /gi=1685101 /ug=Hs.9908 /len=819	39165_at
NCOA3 (nuclear receptor coactivator 3	AF012108	Hs.225977	NM_006534	20q12	Cluster Incl. AF012108:Homo sapiens Amplified in Breast Cancer (AIB1) mRNA, complete cds /cds=(200,4462) /gb=AF012108 /gi=2331249 /ug=Hs.225977 /len=6818	33381_at
LEF1 (lymphoid enhancer-binding factor 1)	AL049409	Hs.44865	NM_016269	4q23-q25	Cluster Incl. AL049409:Homo sapiens mRNA; cDNA DKFZp586H0919 (from clone DKFZp586H0919) /cds=UNKNOWN /gb=AL049409 /gi=4500194 /ug=Hs.44865 /len=1419	36021_at
SIAT1 (sialyltransferase 1 (beta-galactoside alpha-2,6-sialytransferase))	X62822	Hs.2554	NM_003032	3q27-q28	Cluster Incl. X62822:H.sapiens gene encoding beta-galactoside alpha-2,6-sialyltransferase /cds=(310,1530)	41352_at

					/gb=X62822 /gi=29433 /ug=Hs.2554 /len=2699	
BLNK (B-cell linker)	AF068180	Hs.167746		10q23.2-q23.33	Cluster Incl. AF068180:Homo sapiens B cell linker protein BLNK mRNA, alternatively spliced, complete cds /cds=(153,1523) /gb=AF068180 /gi=3406748 /ug=Hs.167746 /len=1790	38242_at
SIAT1 (sialyltransferase 1 (beta-galactoside alpha-2,6-sialytransferase))	W30677	Hs.2554	NM_003032	3q27-q28	Cluster Incl. W30677:zb75h10.r1 Homo sapiens cDNA, 5 end /clone=IMAGE-309475 /clone_end=5 /gb=W30677 /gi=1311730 /ug=Hs.5019 /len=614	34871_at
PSCD1 (pleckstrin homology, Sec7 and coiled/coil domains 1(cytohesin 1)	M85169	Hs.1050	NM_004762	17q25	Cluster Incl. M85169:Human homologue of yeast sec7 mRNA, complete cds /cds=(69,1265) /gb=M85169 /gi=338001 /ug=Hs.1050 /len=3301	38666_at
LOC54103(hypothetical protein)	AL079277	Hs.12969		7	Cluster Incl. AL079277:Homo sapiens mRNA full length insert cDNA clone EUROIMAGE 293605 /cds=(0,806)	41710_at

					/gb=AL079277 /gi=5102581 /ug=Hs.12969 /len=1414	
	U23852				Cluster Incl. U23852:Human T-lymphocyte specific protein tyrosine kinase p56lck (lck) abberant mRNA, complete cds /cds=(59,1150) /gb=U23852 /gi=775207 /ug=Hs.1765 /len=2129	33238_at
KIAA0922(KIAA0922 protein)	AB023139	Hs.37892	NM_015196	4	Cluster Incl. AB023139:Homo sapiens mRNA for KIAA0922 protein, partial cds /cds=(0,2372) /gb=AB023139 /gi=4589475 /ug=Hs.37892 /len=2505	39929_at
	AL049471				Cluster Incl. AL049471:Homo sapiens mRNA; cDNA DKFZp586N012 (from clone DKFZp586N012) /cds=UNKNOWN /gb=AL049471 /gi=4500266 /ug=Hs.12702 /len=2905	41690_at
ISG20 (interferon stimulated gene (20kD))	U88964	Hs.183487	NM_002201	15q26	Cluster Incl. U88964:Human HEM45 mRNA, complete cds /cds=(37,582)	33304_at

					/gb=U88964 /gi=2062679 /ug=Hs.183487 /len=701	
SYNE-2(synaptic nuclei expressed gene 2)	AL080133	Hs.57749	NM_015180	14	Cluster Incl. AL080133:Homo sapiens mRNA; cDNA DKFZp434G173 (from clone DKFZp434G173) /cds=(122,3400) /gb=AL080133 /gi=5262573 /ug=Hs.57749 /len=4307	41815_at
SETBP1 (SET binding protein 1)	AB022660	Hs.151717	NM_015559	18q21.1	Cluster Incl. AB022660:Homo sapiens mRNA for SET-binding protein (SEB), complete cds /cds=(5,4633) /gb=AB022660 /gi=5478317 /ug=Hs.151717 /len=5744	34990_at
FLJ10140(hypothetical protein FLJ10140)]	AL031588	Hs.250671	NM_018006	22	Cluster Incl. AL031588:dJ1163J1.1 (ortholog of mouse transmembrane receptor Celsr1 (KIAA0279 LIKE EGF-like domain containing protein similar to rat MEG /cds=(0,4433) /gb=AL031588 /gi=4007108 /ug=Hs.123043 /len=6438	41660_at

SIT(SHP2 interacting transmembrane adaptor)	AJ010059:	Hs.88012	NM_014450	9	Cluster Incl. AJ010059:Homo sapiens SIT protein /cds=(87,677) /gb=AJ010059 /gi=4688891 /ug=Hs.88012 /len=1232	40723_at
SYNE-1B(synaptic nuclear envelope 1)	AB018339	Hs.8182		6	Cluster Incl. AB018339:Homo sapiens mRNA for KIAA0796 protein, partial cds /cds=(0,3243) /gb=AB018339 /gi=3882312 /ug=Hs.8182 /len=3900	38113_at
HLA-DOB (major histocompatibility complex, class II, DO beta)	X03066:	Hs.1802	NM_002120	6p21.3	Cluster Incl. X03066:Human mRNA for HLA-D class II antigen DO beta chain /cds=(56,877) /gb=X03066 /gi=32271 /ug=Hs.1802 /len=1322	38570_at
POU2AF1 (POU domain, class 2, associating factor 1)	Z49194	Hs.2407	NM_006235	11q23.1	Cluster Incl. Z49194:H.sapiens mRNA for oct-binding factor /cds=(523,1293) /gb=Z49194 /gi=974830 /ug=Hs.2407 /len=3301	36239_at
EZH1 (enhancer of zeste (Drosophila) homolog 1)	AB002386	Hs.194669		17q21.1-q21.3	Cluster Incl. AB002386:Human mRNA for KIAA0388 gene, complete cds /cds=(100,2343) /gb=AB002386	32259_at

					/gi=2224716 /ug=Hs.194669 /len=4606	
SP140(nuclear body protein Sp140)	U36500	Hs.309943	NM_007237	2	Cluster Incl. U36500:Human lymphoid-specific SP100 homolog (LYSP100-B) mRNA, complete cds /cds=(116,2764) /gb=U36500 /gi=1173653 /ug=Hs.85283 /len=3252	40700_at
MTMR1 (myotubularin related protein 1)	AJ224979	Hs.23200		xq28	Cluster Incl. AJ224979:Homo sapiens mRNA for MTMR1 protein /cds=(0,1990) /gb=AJ224979 /gi=4128155 /ug=Hs.23200 /len=2582	34654_at
KIAA0640(SWAP-70 protein	AB014540	Hs.153026		11	Cluster Incl. AB014540:Homo sapiens mRNA for KIAA0640 protein, partial cds /cds=(0,1812) /gb=AB014540 /gi=3327093 /ug=Hs.153026 /len=4824	31869_at
CCR7 (chemokine (C-C motif) receptor 7)	L31584	Hs.1652	NM_001838	17q12-q21.2	L31584 /FEATURE=exon /DEFINITION=HUMEBI103 Human G protein-coupled receptor (EBI 1) gene	1097_s_at

					exon 3, complete cds	
SCAP1 (src family associated phosphoprotein 1)	Y11215	Hs.19126	NM_003726	17q21.3	Cluster Incl. Y11215:Homo sapiens mRNA for SKAP55 protein /cds=(70,1149) /gb=Y11215 /gi=2252495 /ug=Hs.19126 /len=1524	38862_at
IL24 (interleukin 24)	AA214546	Hs.315463	NM_006850	1q32	Cluster Incl. AA214546:zr92c03.s1 Homo sapiens cDNA, 3 end /clone=IMAGE-683140 /clone_end=3 /gb=AA214546 /gi=1813171 /ug=Hs.66576 /len=516	41847_at
ABLIM (actin binding LIM protein)	D31883	Hs.158203	NM_002313	10q25	Cluster Incl. D31883:Human mRNA for KIAA0059 gene, complete cds /cds=(221,1609) /gb=D31883 /gi=505093 /ug=Hs.158203 /len=6754	40155_at
JAK1 (Janus kinase 1 (a protein tyrosine kinase))	AL039831	Hs.50651	NM_002227	1p32.3-p31.3	Cluster Incl. AL039831:DKFZp434D1112_s1 Homo sapiens cDNA, 3 end /clone=DKFZp434D1112 /clone_end=3 /gb=AL039831 /gi=5866713 /ug=Hs.50651	34877_at

					/len=579	
TLK1 (tousled-like kinase 1)	D50927	Hs.18895	NM_012290	8p22-p12	Cluster Incl. D50927:Human mRNA for KIAA0137 gene, complete cds /cds=(1088,2737) /gb=D50927 /gi=1469196 /ug=Hs.18895 /len=4454	32219_at
MGEA5 (meningioma expressed antigen 5 (hyaluronidase))	AB014579	Hs.5734	NM_012215	10q24.1-q24.3	Cluster Incl. AB014579:Homo sapiens mRNA for KIAA0679 protein, partial cds /cds=(0,2303) /gb=AB014579 /gi=3327171 /ug=Hs.5734 /len=4303	35317_at
AIF1 (allograft inflammatory factor 1)	Y14768	Hs.76364	NM_001623	6p21.3	Cluster Incl. Y14768:Homo sapiens DNA, cosmid clones TN62 and TN82 /cds=(10,744) /gb=Y14768 /gi=3805800 /ug=Hs.890 /len=896	40729_s_at
KIAA0430(KIAA0430 gene product)	AB007890			16	Cluster Incl. AB007890:Homo sapiens KIAA0430 mRNA, complete cds /cds=(0,3172) /gb=AB007890 /gi=2887438	31936_s_at

					/ug=Hs.166163 /len=6011	
LCK (lymphocyte-specific protein tyrosine kinase)	M36881	Hs.1765	NM_005356	1p35-p34.3	M36881 /FEATURE=mRNA /DEFINITION=HUMLCKAA Human lymphocyte-specific protein tyrosine kinase (lck) mRNA, complete cds	2059_s_at
GPR18 (G protein-coupled receptor 18)	L42324	Hs.88269		13q32	L42324 /FEATURE=cds /DEFINITION=HUMFRCG Homo sapiens (clone GPCR W) G protein-linked receptor gene (GPCR) gene, 5 end of cds	253_g_at
TC21(oncogene TC21)	AI365215	Hs.206097	NM_012250	11	Cluster Incl. AI365215:qz41a06.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-2029426 /clone_end=3 /gb=AI365215 /gi=4124904 /ug=Hs.206097 /len=918	32827_at
	AI434146				Cluster Incl. AI434146:ti36g07.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-2132604 /clone_end=3 /gb=AI434146 /gi=4294137 /ug=Hs.164284 /len=299	36403_s_at

AKAP11 (A kinase (PRKA) anchor protein 11)	AB014529	Hs.232076	NM_016248	13q12.2-13q14.3	Cluster Incl. AB014529:Homo sapiens mRNA for KIAA0629 protein, partial cds /cds=(0,1840) /gb=AB014529 /gi=3327071 /ug=Hs.232076 /len=5883	34657_at
IL4R (interleukin 4 receptor)	X52425	Hs.75545	NM_000418	16p11.2-12.1	X52425 /FEATURE=mRNA /DEFINITION=HSIL4R Human IL-4-R mRNA for the interleukin 4 receptor	404_at
ARHH (ras homolog gene family, member H)	Z35227	Hs.109918	NM_004310	4p13	Cluster Incl. Z35227:H.sapiens TTF mRNA for small G protein /cds=(579,1154) /gb=Z35227 /gi=609016 /ug=Hs.109918 /len=1427	37416_at
E2F5 (E2F transcription factor 5, p130-binding)	U31556	Hs.2331	NM_001951	8p22-q21.3	U31556 /FEATURE= /DEFINITION=HSU31556 Human transcription factor E2F-5 mRNA, complete cds	1044_s_at
TOSO(regulator of Fas-induced apoptosis)]	AF057557	Hs.58831	NM_005449	1	Cluster Incl. AF057557:Homo sapiens anti-Fas-induced apoptosis (TOSO) mRNA, complete cds /cds=(19,1191)	32967_at

					/gb=AF057557 /gi=3169292 /ug=Hs.238857 /len=1339	
E2F5 (E2F transcription factor 5, p130-binding)	U31556	Hs.2331	NM_001951	8p22-q21.3	Cluster Incl. U31556:Human transcription factor E2F-5 mRNA, complete cds /cds=(38,1075) /gb=U31556 /gi=939728 /ug=Hs.2331 /len=1728	41275_at
PIR121(p53 inducible protein)	L47738	Hs.258503		5	Cluster Incl. L47738:Homo sapiens inducible protein mRNA, complete cds /cds=(1004,1714) /gb=L47738 /gi=1009098 /ug=Hs.80313 /len=2881	37579_at
KIAA0543(KIAA0543 protein)	AB011115	Hs.98507		7	Cluster Incl. AB011115:Homo sapiens mRNA for KIAA0543 protein, partial cds /cds=(0,3336) /gb=AB011115 /gi=3043609 /ug=Hs.98507 /len=6443	41077_at
KIAA0769(KIAA0769 gene product)]	AB018312	Hs.19056	NM_014824	11	Cluster Incl. AB018312:Homo sapiens mRNA for KIAA0769 protein, complete cds /cds=(239,2293) /gb=AB018312	32224_at

					/gi=3882258 /ug=Hs.19056 /len=4326	
	AB023183	Hs.52463	Hs.52463	10	Cluster Incl. AB023183:Homo sapiens mRNA for KIAA0966 protein, complete cds /cds=(166,3564) /gb=AB023183 /gi=4589575 /ug=Hs.52463 /len=4924	36089_at
PRDM2 (PR domain containing 2, with ZNF domain)	D45132	Hs.26719	NM_012231	1p36	D45132 /FEATURE= /DEFINITION=HUMHOXY1 Homo sapiens mRNA for zinc-finger DNA-binding protein, complete cds	316_g_at
BCL2 (B-cell CLL/lymphoma 2)	M14745	Hs.79241	NM_000633	18q21.3	M14745 /FEATURE= /DEFINITION=HUMBCL2C Human bcl-2 mRNA	1909_at
KIAA0239(KIAA0239 protein)	D87076	Hs.9729	NM_015288	5	Cluster Incl. D87076:Human mRNA for KIAA0239 gene, partial cds /cds=(0,1716) /gb=D87076 /gi=1510152 /ug=Hs.9729 /len=5630	38342_at

DKFZP434L243(DKFZP434L243 protein	AL080140	Hs.21695		3	Cluster Incl. AL080140:Homo sapiens mRNA; cDNA DKFZp434L243 (from clone DKFZp434L243) /cds=(0,2137) /gb=AL080140 /gi=5262585 /ug=Hs.21695 /len=2662	34220_at
IFI41 (interferon-induced protein 41, 30kD)	L22342	Hs.241510	NM_004509		Cluster Incl. L22342:Human nuclear phosphoprotein mRNA, complete cds /cds=(0,746) /gb=L22342 /gi=402204 /ug=Hs.38125 /len=835	35718_at
SLC23A1 (solute carrier family 23 (nucleobase transporters), member 1)	D87075	Hs.82042	NM_005116	20p13	Cluster Incl. D87075:Human mRNA for KIAA0238 gene, partial cds /cds=(0,992) /gb=D87075 /gi=1510150 /ug=Hs.82042 /len=5608	38122_at
PPP2R5C (protein phosphatase 2, regulatory subunit B (B56), gamma isoform)	U37352	Hs.171734	NM_002719	3p21	Cluster Incl. U37352:Human protein phosphatase 2A Balpha1 regulatory subunit mRNA, complete cds /cds=(88,1632) /gb=U37352 /gi=1203811 /ug=Hs.171734 /len=4064	40786_at

MSF (MLL septin-like fusion (NOTE: non-standard symbol and name))	AB023208	Hs.181002	NM_006640	17q25	Cluster Incl. AB023208:Homo sapiens mRNA for KIAA0991 protein, complete cds /cds=(732,2000) /gb=AB023208 /gi=4589625 /ug=Hs.181002 /len=3938	41220_at
GG2-1(TNF-induced protein)	AF099935	Hs.17839	NM_014350	5	Cluster Incl. AF099935:Homo sapiens MDC-3.13 isoform 2 mRNA, complete cds /cds=(84,680) /gb=AF099935 /gi=3860092 /ug=Hs.17839 /len=1897	33243_at
CBLB (Cas-Br-M (murine) ectropic retroviral transforming sequence b)	U26710	Hs.3144	NM_004351	3p13-q13.2	Cluster Incl. U26710:Human cbl-b mRNA, complete cds /cds=(322,3270) /gb=U26710 /gi=862406 /ug=Hs.3144 /len=3969	35632_at
E2F5 (E2F transcription factor 5, p130-binding)	U15642	Hs.2331	NM_001951	8p22-q21.3	U15642 /FEATURE= /DEFINITION=HSU15642 Human transcription factor E2F-5 mRNA, complete cds	1639_s_at
GGA2(Golgi-associated, gamma-adaptin ear containing, ARF-binding protein 2)	AB029003	Hs.155546	NM_015044	16	Cluster Incl. AB029003:Homo sapiens mRNA for KIAA1080 protein, partial cds	40278_at

containing, ARF-binding protein 2)					/cds=(0,1554) /gb=AB029003 /gi=5689496 /ug=Hs.155546 /len=4791	
KIAA0240(KIAA0240 protein)]	D87077	Hs.196275		6	Cluster Incl. D87077:Human mRNA for KIAA0240 gene, partial cds /cds=(0,2953) /gb=D87077 /gi=1510154 /ug=Hs.196275 /len=6060	38892_at
KIAA0542(KIAA0542 gene product)	AB011114	Hs.62209		22	Cluster Incl. AB011114:Homo sapiens mRNA for KIAA0542 protein, complete cds /cds=(393,3299) /gb=AB011114 /gi=3043607 /ug=Hs.62209 /len=5280	36545_s_at
DKFZP586F2423(hypothetical protein DKFZp586F2423)	AL080209:	Hs.13659		7	Cluster Incl. AL080209:Homo sapiens mRNA; cDNA DKFZp586F2423 (from clone DKFZp586F2423) /cds=UNKNOWN /gb=AL080209 /gi=5262698 /ug=Hs.13659 /len=4241	39692_at
GPR18 (G protein-coupled receptor 18)	L42324	Hs.88269		13q32	L42324 /FEATURE=cds /DEFINITION=HUMFRCG Homo sapiens (clone GPCR W) G protein-linked receptor	252_at

					gene (GPCR) gene, 5 end of cds	
LIG1 (ligase I, DNA, ATP-dependent)	AL039458	Hs.4193		3p14	Cluster Incl. AL039458:DKFZp434N0910_s1 Homo sapiens cDNA, 3 end /clone=DKFZp434N0910 /clone_end=3 /gb=AL039458 /gi=5408506 /ug=Hs.4193 /len=849	34800_at
KIAA0136(DNA segment, Chr 16, Johns Hopkins University 32, expressed)	D50926	Hs.70359		21q22.13	Cluster Incl. D50926:Human mRNA for KIAA0136 gene, partial cds /cds=(0,2854) /gb=D50926 /gi=1469194 /ug=Hs.70359 /len=4197	36845_at
SGF3G (interferon-stimulated transcription factor 3, gamma (48kD))	M87503	Hs.1706	NM_006084	14q11.2	Cluster Incl. M87503:Human IFN-responsive transcription factor subunit mRNA, complete cds /cds=(34,1215) /gb=M87503 /gi=184652 /ug=Hs.1706 /len=1584	38517_at

KIAA0441(KIAA0441 gene product)	AB007901			6	Cluster Incl. AB007901:Homo sapiens KIAA0441 mRNA, complete cds /cds=(168,2261) /gb=AB007901 /gi=2662162 /ug=Hs.32511 /len=5597	39658_at
P2Y10(putative purinergic receptor)	AF000545	Hs.296433	NM_014499	X	AF000545 /FEATURE=cds /DEFINITION=HSAF000545 Homo sapiens putative purinergic receptor P2Y10 gene, complete cds	358_at
PPP3CC (protein phosphatase 3 (formerly 2B), catalytic subunit, gamma isoform (calcineurin A gamma))	S46622	Hs.75206	NM_005605	8	Cluster Incl. S46622:calcineurin A catalytic subunit [human, testis, mRNA, 2134 nt] /cds=(286,1794) /gb=S46622 /gi=258000 /ug=Hs.75206 /len=2134	32541_at
MGC12335(hypothetical protein MGC12335)	AL022724	Hs.97411	NM_032744	6	Cluster Incl. AL022724:Human DNA sequence from clone 413H6 on chromosome 6p22.3-24.3. Contains a hamster Androgen-dependent Expressed Protein like protein gene, ESTs and GSSs /cds=(94,861) /gb=AL022724 /gi=4468306	34043_at

					/ug=Hs.97411 /len=1037	
SP100 (nuclear antigen Sp100)	M60618	Hs.77617	NM_003113	2q37.1	Cluster Incl. M60618:Human nuclear autoantigen (SP-100) mRNA, complete cds /cds=(31,1473) /gb=M60618 /gi=178688 /ug=Hs.77617 /len=1879	37352_at
KIAA0746(KIAA0746 protein)]	AB018289	Hs.49500		4	Cluster Incl. AB018289:Homo sapiens mRNA for KIAA0746 protein, partial cds /cds=(0,3091) /gb=AB018289 /gi=3882212 /ug=Hs.49500 /len=4086	41585_at
RBL2 (retinoblastoma-like 2 (p130))	X76061	Hs.79362	NM_005611	16q12.2	Cluster Incl. X76061:H.sapiens p130 mRNA for 130K protein /cds=(69,3488) /gb=X76061 /gi=416030 /ug=Hs.79362 /len=4835	32597_at
APOC4 (apolipoprotein C-IV)	U32576	Hs.110675	NM_001646	19q13.2	Cluster Incl. U32576:Human apolipoprotein apoC-IV (APOC4) gene, complete cds /cds=(40,423) /gb=U32576	34454_r_at

					/gi=975892 /ug=Hs.110675 /len=613	
STK10 (serine/threonine kinase 10)	AB015718	Hs.16134	NM_005990	5q35.1	Cluster Incl. AB015718:Homo sapiens lok mRNA for protein kinase, complete cds /cds=(50,2956) /gb=AB015718 /gi=4001687 /ug=Hs.16134 /len=4221	40420_at
MAP3K5 (mitogen-activated protein kinase kinase kinase 5)	U67156	Hs.151988	NM_005923	6q22.33	U67156 /FEATURE= /DEFINITION=HSU67156 Human mitogen-activated kinase kinase kinase 5 (MAPKKK5) mRNA, complete cds	1327_s_at
PRKCB1 (protein kinase C, beta 1)	X07109	Hs.77202	NM_002738	16p11.2	X07109 /FEATURE=cds /DEFINITION=HSPKCB2A Human mRNA for protein kinase C (PKC) type beta II /NOTE=replacement of probe set 1216_at	160029_at
BIRC3 (baculoviral IAP repeat-containing 3)	U45878	Hs.127799	NM_001165	11q22	U45878 /FEATURE= /DEFINITION=HSU45878 Human inhibitor of apoptosis protein 1 mRNA, complete cds	1717_s_at

SLC7A6 (solute carrier family 7 (cationic amino acid transporter, y+ system), member 6)	D87432	Hs.10315	NM_003983	16q22.1-q22.3	Cluster Incl. D87432:Human mRNA for KIAA0245 gene, complete cds /cds=(261,1808) /gb=D87432 /gi=1665758 /ug=Hs.10315 /len=6296	39533_at
KIAA1093(KIAA1093 protein)	AB029016	Hs.117333		22	Cluster Incl. AB029016:Homo sapiens mRNA for KIAA1093 protein, partial cds /cds=(0,3613) /gb=AB029016 /gi=5689522 /ug=Hs.117333 /len=4159	37487_at

Table 10:

UCL/HGNC/HUGO Human Gene Nomenclature Database Symbol	GenBank Accession No.	UniGene Cluster	RefSeq	Chromosomal Location	Description Unigene Build #95	Gene Name
GATA2 (GATA-binding protein 2)	M77810	Hs.334695	NM_002050	3q21	M77810 /FEATURE= /DEFINITION=HUMGATA2A Human transcription factor GATA-2 (GATA-2) mRNA, complete cds	1072_g_at
MYL6 (myosin, light polypeptide 6, alkali, smooth muscle and non-muscle)	M22919	Hs.77385	NM_021019	12	Cluster Incl. M22919:Human nonmuscle/smooth muscle alkali myosin light chain gene, complete cds /cds=(42,353) /gb=M22919 /gi=189016 /ug=Hs.77385 /len=1259	33994_g_at
PRG2 (proteoglycan 2, bone marrow (natural killer cell activator, eosinophil granule major basic protein))	Z26248	Hs.99962	NM_002728	11q12	Cluster Incl. Z26248:H.sapiens mRNA for eosinophil granule major basic protein /cds=(857,1525) /gb=Z26248 /gi=940510 /ug=Hs.99962 /len=1637	39179_at

					/ug=Hs.99962 /len=1637	
CLC (Charot-Leyden crystal protein)	L01664	Hs.132004	NM_013246	11q13.3	Cluster Incl. L01664:Human eosinophil Charcot-Leyden crystal (CLC) protein (lysophospholipase) mRNA, complete cds /cds=(33,461) /gb=L01664 /gi=187273 /ug=Hs.889 /len=586	36809_at
ST7 (suppression of tumorigenicity 7)	W02490	Hs.301974	NM_013437	8q22.2-q23.1	Cluster Incl. W02490:za48b02.r1 Homo sapiens cDNA, 5 end /clone=IMAGE-295755 /clone_end=5 /gb=W02490 /gi=1274488 /ug=Hs.5814 /len=623	40039_g_at
TTF2 (transcription termination factor, RNA polymerase II)	AF073771	Hs.142157	NM_003594	1p22	Cluster Incl. AF073771:Homo sapiens RNA polymerase II termination factor mRNA, complete cds /cds=(20,3508) /gb=AF073771 /gi=3702845 /ug=Hs.142157 /len=3591	37870_at
TALDO1 (transaldolase 1)	AF010400	Hs.77290	NM_006755	11p15.5-p15.4	Cluster Incl. AF010400:untitled /cds=(50,1063) /gb=AF010400	37311_at

					/gi=2612878 /ug=Hs.77290 /len=1242	
PGD (phosphogluconate dehydrogenase)	U30255	Hs.75888	NM_002631	1p36.3-p36.13	Cluster Incl. U30255:Human phosphogluconate dehydrogenase (hPGDH) gene, complete cds /cds=(6,1457) /gb=U30255 /gi=984324 /ug=Hs.75888 /len=1536	36963_at
GAPD (glyceraldehyde-3-phosphate dehydrogenase)	U34995	Hs.169476	NM_002046	12p13	Cluster Incl. U34995:Human normal keratinocyte substraction library mRNA, clone H22a, complete sequence /cds=UNKNOWN /gb=U34995 /gi=1497857 /ug=Hs.195188 /len=1626	35905_s_at
RNASE2 (ribonuclease, RNase A family, 2 (liver, eosinophil-derived neurotoxin))	X55988	Hs.728	NM_002934	14q24-q31	Cluster Incl. X55988:Human EDN mRNA for eosinophil derived neurotoxin /cds=(71,556) /gb=X55988 /gi=31088 /ug=Hs.728 /len=735	36766_at
ICA1 (islet cell autoantigen 1 (69kD))	U38260	Hs.167927	NM_004968	7p22	Cluster Incl. U38260:Human islet cell autoantigen ICAp69 mRNA, complete cds /cds=(169,942) /gb=U38260 /gi=1675205	32634_s_at

					/ug=Hs.167927 /len=1415	
M6PR (mannose-6-phosphate receptor (cation dependent))	X56253	Hs.75709	NM_002355	12p13	Cluster Incl. X56253:Human MPR46 gene for 46kd mannose 6-phosphate receptor /cds=(168,1001) /gb=X56253 /gi=34727 /ug=Hs.75709 /len=2455	32547_at
GCDH (glutaryl-Coenzyme A dehydrogenase)	AD000092	Hs.184141	NM_000159	19p13.2	AD000092 /FEATURE=cds#4 /DEFINITION=CH19HHR23 Homo sapiens DNA from chromosome 19p13.2 cosmids R31240, R30272 and R28549 containing the EKLF, GCDH, CRTC, and RAD23A genes, genomic sequence	1749_at
ACTB (actin, beta)	X00351	Hs.288061	NM_001101	7p15-p12	Homo sapiens /REF=X00351 /DEF=Human mRNA for beta-actin /LEN=1761 (_5, _M, _3 represent transcript regions 5 prime, Middle, and 3 prime respectively)	AFFX-HSAC0
PMS2L11 (postmeiotic segregation increased 2-like 11)	U38980	Hs.306174		7q	U38980 /FEATURE= /DEFINITION=U38980 Human PMS2	179_at

like 11)					related (hPMSR6) mRNA, complete cds	
NCF4 (neutrophil cytosolic factor 4 (40kD))	AL008637	Hs.196352	NM_000631	22q13.1	Cluster Incl. AL008637:Human DNA sequence from clone 833B7 on chromosome 22q12.3-13.2 Contains genes for NCF4 (P40PHOX) protein,cytokine receptor common beta chain precursor CSF2RB (partial), ESTs, CA repeat, STS, GSS /cds=(629,1648) /gb=AL008637 /gi=3136	38894_g_at
CSF2RB (colony stimulating factor 2 receptor, beta, low-affinity (granulocyte-macrophage))	AL008637	Hs.285401	NM_000395	22q13.1	Cluster Incl. AL008637:Human DNA sequence from clone 833B7 on chromosome 22q12.3-13.2 Contains genes for NCF4 (P40PHOX) protein,cytokine receptor common beta chain precursor CSF2RB (partial), ESTs, CA repeat, STS, GSS /cds=(629,1648) /gb=AL008637 /gi=3136	
TMSB4X (thymosin, beta 4, X chromosome)	M17733	Hs.75968	NM_021109	xq21.3-q22	Cluster Incl. M17733:Human thymosin beta-4 mRNA, complete cds /cds=(77,211)	31557_at

					/gb=M17733 /gi=339688 /ug=Hs.75968 /len=556	
MTX1 (metaxin 1)	U46920	Hs.247551	NM_002455	1q21	Cluster Incl. U46920:Human metaxin (MTX) gene, complete cds /cds=(0,953) /gb=U46920 /gi=1326107 /ug=Hs.181246 /len=1065	40890_at
	AA524802				Cluster Incl. AA524802:nh33h11.s1 Homo sapiens cDNA /clone=IMAGE-954213 /gb=AA524802 /gi=2265730 /ug=Hs.203907 /len=500	32877_i_at
CFL1 (cofilin 1 (non-muscle))	X95404	Hs.180370	NM_005507	11q13	Cluster Incl. X95404:H.sapiens mRNA for non-muscle type cofilin /cds=(51,551) /gb=X95404 /gi=1177470 /ug=Hs.180370 /len=1059	33659_at
DEFA1 (defensin, alpha 1, myeloid-related sequence)	AL036554	Hs.274463	NM_004084	8p23.2-p23.1	Cluster Incl. AL036554:DKFZp564J2262_r1 Homo sapiens cDNA, 5 end /clone=DKFZp564J2262 /clone_end=5	31793_at

					/gb=AL036554 /gi=5927801 /ug=Hs.1379 /len=517	
S100A8 (*S100 calcium-binding protein A8 (calgranulin A)*)	AI126134	Hs.100000	NM_002964	1q21	Cluster Incl. AI126134:qd77c05.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-1735496 /clone_end=3 /gb=AI126134 /gi=3594648 /ug=Hs.100000 /len=446	41096_at
	AC005764				Cluster Incl. AC005764:Homo sapiens chromosome 19, cosmid R31343 /cds=(0,1262) /gb=AC005764 /gi=3694626 /ug=Hs.126496 /len=1263	35512_at
DAPK2 (*death-associated protein kinase 2*)	AF052941	Hs.129208	NM_014326	15	Cluster Incl. AF052941:Homo sapiens DAP-kinase related protein 1 mRNA, complete cds /cds=(31,1143) /gb=AF052941 /gi=3560542 /ug=Hs.129208 /len=1742	34912_at
DEFA3 (*defensin, alpha 3, neutrophil-specific*)	L12691	Hs.294176	NM_005217	8pter-p23.3	Cluster Incl. L12691:Human neutrophil peptide-3 gene, complete cds /cds=(50,334) /gb=L12691 /gi=292364	31506_s_at

					/ug=Hs.178741 /len=452	
MTHFS (5,10-methenyltetrahydrofolate synthetase (5-formyltetrahydrofolate cyclo-ligase))	L38928	Hs.118131	NM_006441	15	Cluster Incl. L38928:Homo sapiens 5,10-methenyltetrahydrofolate synthetase mRNA, complete cds /cds=(13,624) /gb=L38928 /gi=886296 /ug=Hs.118131 /len=857	39064_at
B2M (beta-2-microglobulin)	V00567	Hs.75415	NM_004048	15q21-q22.2	V00567 /FEATURE=cds /DEFINITION=HSMGLO Human messenger RNA fragment for the beta-2 microglobulin	428_s_at
ACTB (actin, beta)	X00351	Hs.288061	NM_001101	7p15-p12	Homo sapiens /REF=X00351 /DEF=Human mRNA for beta-actin /LEN=1761 (_5, _M, _3 represent transcript regions 5 prime, Middle, and 3 prime respectively)	AFFX-HSAC0
ADAM8 (a disintegrin and metalloproteinase domain 8)	D26579	Hs.86947	NM_001109	10q26.3	Cluster Incl. D26579:Homo sapiens mRNA for transmembrane protein, complete cds /cds=(9,2483) /gb=D26579 /gi=1864004	40712_at

					/ug=Hs.86947 /len=3236	
CHRNE (cholinergic receptor, nicotinic, epsilon polypeptide)	X66403	Hs.278295	NM_000080	17p13-p12	Cluster Incl. X66403:H.sapiens mRNA for acetylcholine receptor (epsilon subunit) /cds=(11,1492) /gb=X66403 /gi=560152 /ug=Hs.112028 /len=2457	39834_at
BRCA1 (breast cancer 1, early onset)	L78833	Hs.194143	NM_007294	17q21	L78833 /FEATURE=exon#36 /DEFINITION=HUMBRCA1 Human BRCA1, Rho7 and vatI genes, complete cds, and ipf35 gene, partial cds	605_at
OAZ1 (ornithine decarboxylase antizyme 1)	D78361	Hs.125078		19p13.3	D78361 /FEATURE= /DEFINITION=HUMODAZ Human mRNA for ornithine decarboxylase antizyme, ORF 1 and ORF 2	1315_at
DOC2B (double C2-like domains, beta)	D70830	Hs.54402	NM_003585	17	Cluster Incl. D70830:Homo sapiens mRNA for Doc2 beta, complete cds /cds=(160,1398) /gb=D70830 /gi=1235721 /ug=Hs.54402 /len=2043	32422_at

LIG3 (ligase III, DNA, ATP-dependent)	X84740	Hs.100299	NM_002311	17q11.2-q12	Cluster Incl. X84740:H.sapiens mRNA for DNA ligase III /cds=(333,3101) /gb=X84740 /gi=860962 /ug=Hs.100299 /len=3400	41099_at
S100A9 (S100 calcium-binding protein A9 (calgranulin B))	W72424	Hs.112405	NM_002965	1q21	Cluster Incl. W72424:zd66a09.s1 Homo sapiens cDNA, 3 end /clone=IMAGE-345592 /clone_end=3 /gb=W72424 /gi=1382379 /ug=Hs.112405 /len=604	41471_at
MPB1 (MYC promoter-binding protein 1)	M55914	Hs.284127	NM_005945	1pter-p35	M55914 /FEATURE= /DEFINITION=HUMCMYCQ Human c-myc binding protein (MBP-1) mRNA, complete cds	2035_s_at
MAD (MAX dimerization protein)	L06895	Hs.109012	NM_002357	2p13-p12	L06895 /FEATURE= /DEFINITION=HUMMAD Homo sapiens antagonizer of myc transcriptional activity (Mad) mRNA, complete cds	1774_at

GAPD (glyceraldehyde-3-phosphate dehydrogenase)	M33197	Hs.169476	NM_002046	12p13	Homo sapiens /REF=M33197 /DEF=Human glyceraldehyde-3-phosphate dehydrogenase (GAPDH) mRNA, complete cds /LEN=1268 (_5, _M, _3 represent transcript regions 5 prime, Middle, and 3 prime respectively)	AFFX-HUMGA
RAB31 (RAB31, member RAS oncogene family)	U59877	Hs.223025	NM_006868	18p11.3	Cluster Incl. U59877:Human low-Mr GTP-binding protein (RAB31) mRNA, complete cds /cds=(60,644) /gb=U59877 /gi=1388194 /ug=Hs.223025 /len=907	33371_s_at
ABCC5 (ATP-binding cassette, sub-family C (CFTR/MRP), member 5)	U83661	Hs.108660	NM_005688	3q27	U83661 /FEATURE= /DEFINITION=HSU83661 Homo sapiens multidrug resistance protein 5 (MRP5) mRNA, complete cds	1933_g_at
PABPC2 (poly(A)-binding protein, cytoplasmic 2)	Z48501	Hs.172182			Cluster Incl. Z48501:H.sapiens mRNA for polyadenylate binding protein II /cds=(0,1568) /gb=Z48501 /gi=693936 /ug=Hs.172182 /len=1569	31951_s_at

HBB (hemoglobin, beta)	L48215	Hs.155376	NM_000518	11p15.5	Cluster Incl. L48215:Homo sapiens beta-globin (HBB) gene, with a to c allele 28 bp 5 to exon 1, (J00179 bases 61971-63802) /cds=(50,493) /gb=L48215 /gi=1066772 /ug=Hs.155376 /len=626	32052_at
P8(nuclear proten 1)	W47047	Hs.8603	NM_012385	16	Cluster Incl. W47047:zc38g10.r1 Homo sapiens cDNA, 5 end /clone=IMAGE-324642 /clone_end=5 /gb=W47047 /gi=1331686 /ug=Hs.166194 /len=441	36423_at
	M64936				M64936 /FEATURE= /DEFINITION=HUMRIRT Homo sapiens retinoic acid-inducible endogenous retroviral DNA	1090_f_at
	M64936				Cluster Incl. M64936:Homo sapiens retinoic acid-inducible endogenous retroviral DNA /cds=UNKNOWN /gb=M64936 /gi=337422 /ug=Hs.55322 /len=3307	36727_at

Human BRCA2 region	U50534	Hs.181304	NM_023037	13	U50534 /FEATURE= /DEFINITION=HSU50534 Human BRCA2 region, mRNA sequence CG003	1529_at
	AL049675				Cluster Incl. AL049675:Human gene from PAC 886K2, chromosome 1 /cds=UNKNOWN /gb=AL049675 /gi=4678768 /ug=Hs.15535 /len=1074	32048_at
CAMK2B (calcium/calmodulin-dependent protein kinase (CaM kinase) II beta)	AF112471	Hs.4884	NM_001220	7p14.3-p14.1	Cluster Incl. AF112471:Homo sapiens calcium/calmodulin-dependent protein kinase II beta subunit mRNA, alternatively spliced, complete cds /cds=(46,1599) /gb=AF112471 /gi=4139267 /ug=Hs.4884 /len=1750	34847_s_at
HBB (hemoglobin, beta)	M25079	Hs.155376	NM_000518	11p15.5	Cluster Incl. M25079:Human sickle cell beta-globin mRNA, complete cds /cds=(0,443) /gb=M25079 /gi=179408 /ug=Hs.234764 /len=468	31687_f_at

HSPC022(HSPC022 protein)	W68830	Hs.301175	NM_014029	22	Cluster Incl. W68830:zd37g06.r1 Homo sapiens cDNA, 5 end /clone=IMAGE-342874 /clone_end=5 /gb=W68830 /gi=1377739 /ug=Hs.173466 /len=614	32736_at
B2M (beta-2-microglobulin)	AB021288	Hs.75415	NM_004048	15q21-q22.2	Cluster Incl. AB021288:Homo sapiens mRNA for beta 2-microglobulin, complete cds /cds=(13,372) /gb=AB021288 /gi=4038732 /ug=Hs.75415 /len=925	34644_at
HMG17 (high-mobility group (nonhistone chromosomal) protein 17)	X13546	Hs.181163	NM_005517	1p36.1	Cluster Incl. X13546:Human HMG-17 gene for non-histone chromosomal protein HMG-17 /cds=(107,379) /gb=X13546 /gi=32328 /ug=Hs.181163 /len=1198	41231_f_at
GDF1 (growth differentiation factor 1)	M62302	Hs.336964	NM_001492	19p12	M62302 /FEATURE= /DEFINITION=HUMGDF1 Human growth/differentiation factor 1 (GDF-1) mRNA, complete cds	887_at
EPX (eosinophil peroxidase)	X14346	Hs.46295	NM_000502	17q23.1	Cluster Incl. X14346:Human mRNA for eosinophil peroxidase /cds=(0,2108)	34587_at

					/gb=X14346 /gi=31182 /ug=Hs.46295 /len=2558	
VCL (vinculin)	M33308	Hs.75350	NM_003373	10q22.1-q23	Cluster Incl. M33308:Human vinculin mRNA, complete cds /cds=(50,3250) /gb=M33308 /gi=340236 /ug=Hs.75350 /len=5102	36601_at
CLIC2 (chloride intracellular channel 2)	Y12696	Hs.54570	NM_001289	xq28	Cluster Incl. Y12696:H.sapiens mRNA homologous to the p64 bovine chloride channel peptide /cds=(221,952) /gb=Y12696 /gi=2584784 /ug=Hs.54570 /len=1219	40013_at
RPL41 (ribosomal protein L41)	Z12962	Hs.324406	NM_021104	12q	Homo sapiens /REF=Z12962 /DEF=Cluster Incl. :H.sapiens mRNA for homologue to yeast ribosomal protein L41 /cds=(83,160) /gb= /gi=36135 /ug=Hs.108124 /len=468 /LEN=468	32466_at
TNA (tetranectin (plasminogen-binding protein))	X64559	Hs.65424	NM_003278	3p22-p21.3	Cluster Incl. X64559:H.sapiens mRNA for tetranectin /cds=(93,701) /gb=X64559	36569_at

					/gi=37408 /ug=Hs.65424 /len=848	
MYC (v-myc avian myelocytomatosis viral oncogene homolog)	M13929	Hs.79070	NM_002467	8q24.12-q24.13	M13929 /FEATURE=mRNA /DEFINITION=HUMMYCPOA Human c-myc-P64 mRNA, initiating from promoter P0, (HLmyc2.5) partial cds	1827_s_at
FCGR2A (Fc fragment of IgG, low affinity IIa, receptor for (CD32))	M31932	Hs.78864	NM_021642	1q23	Cluster Incl. M31932:Human IgG low affinity Fc fragment receptor (FcRIIa) mRNA, complete cds /cds=(7,960) /gb=M31932 /gi=182473 /ug=Hs.78864 /len=2372	37687_i_at
RPS3A (ribosomal protein S3A)	M84711	Hs.77039	NM_001006	4q31.2-q31.3	M84711 /FEATURE= /DEFINITION=HUMFTE1A Human v-fos transformation effector protein (Fte-1), mRNA complete cds	1653_at
	H12458				H12458 /FEATURE= /DEFINITION=H12458 yj12d03.s1 Soares placenta Nb2HP Homo sapiens cDNA clone IMAGE:148517 3 similar to	2090_i_at

					SP:WNT6_MOUSE P22727 WNT-6 PROTEIN ;, mRNA sequence	
IRF5 (interferon regulatory factor 5)	U51127	Hs.334450	NM_002200	7q32	U51127 /FEATURE= /DEFINITION=HSU51127 Human interferon regulatory factor 5 (Humirf5) mRNA, complete cds	477_at
H3F3A (H3 histone, family 3A)	M11353	Hs.181307	NM_002107	1q41	M11353 /FEATURE= /DEFINITION=HUMHISH3C Human H3.3 histone class C mRNA, complete cds	254_at
MAPT (microtubule-associated protein tau)		Hs.101174	NM_005910	17q21.1	Microtubule-Associated Protein Tau, Alt. Splice 5, Exon 4a	331_at
CML1(kidney- and liver-specific gene)	AB013094	Hs.14637	NM_003960	2	Cluster Incl. AB013094:Homo sapiens TSC501 mRNA, complete cds /cds=(168,851) /gb=AB013094 /gi=3721765 /ug=Hs.14637 /len=960	38128_at

HCLS1 (hematopoietic cell-specific Lyn substrate 1)	X16663	Hs.14601	NM_005335	3q13	Cluster Incl. X16663:Human HS1 gene for heamatopoietic lineage cell specific protein /cds=(42,1502) /gb=X16663 /gi=32054 /ug=Hs.14601 /len=1950	31820_at
KCNH2 (potassium voltage-gated channel, subfamily H (eag-related), member 2)	U04270	Hs.188021	NM_000238	7q35-q36	Cluster Incl. U04270:Human putative potassium channel subunit (h-erg) mRNA, complete cds /cds=(183,3662) /gb=U04270 /gi=487737 /ug=Hs.188021 /len=4070	38858_at
FTL (ferritin, light polypeptide)	AL031670	Hs.111334	NM_000146	19q13.3-q13.4	Cluster Incl. AL031670:dJ681N20.2 (ferritin, light polypeptide-like 1) /cds=(200,727) /gb=AL031670 /gi=4469083 /ug=Hs.111334 /len=978	35083_at
ALOX5AP (arachidonate 5-lipoxygenase-activating protein)	AI806222	Hs.100194	NM_001629	13q12	Cluster Incl. AI806222:wf26e10.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-2356746 /clone_end=3 /gb=AI806222 /gi=5392788 /ug=Hs.100194 /len=563	37099_at

MLLT7 (myeloid/lymphoid or mixed-lineage leukemia (trithorax (Drosophila) homolog); translocated to, 7)	Y11284	Hs.239663	NM_005938	xq13.1	Cluster Incl. Y11284:Homo sapiens AFX1 gene, exon 1 (and joined CDS) /cds=(244,1758) /gb=Y11284 /gi=2879783 /ug=Hs.239663 /len=3162	36238_at
RPS4X (*ribosomal protein S4, X-linked*)	M58458	Hs.108124	NM_001007	xq13.1	Cluster Incl. M58458:Human ribosomal protein S4 (RPS4X) isoform mRNA, complete cds /cds=(35,826) /gb=M58458 /gi=337509 /ug=Hs.75344 /len=888	34643_at
TLN1 (talin 1)	AB028950	Hs.18420	NM_006289	9p13	Cluster Incl. AB028950:Homo sapiens mRNA for KIAA1027 protein, partial cds /cds=(0,5088) /gb=AB028950 /gi=5689390 /ug=Hs.18420 /len=5542	32166_at
BPHL (biphenyl hydrolase-like (serine hydrolase; breast epithelial mucin-associated antigen))	X81372	Hs.7298	NM_004332	6p25	Cluster Incl. X81372:H.sapiens mRNA for biphenyl hydrolase-related protein /cds=(212,1036) /gb=X81372 /gi=984662 /ug=Hs.184552 /len=1508	40912_s_at
RPA2 (replication protein A2 (32kD))	J05249	Hs.79411	NM_002946	1p35	J05249 /FEATURE= /DEFINITION=HUMREPA Human	1119_at

					replication protein A 32-kDa subunit mRNA, complete cds	
NCF4 (neutrophil cytosolic factor 4 (40kD))	AL008637	Hs.196352	NM_000631	22q13.1	Cluster Incl. AL008637:Human DNA sequence from clone 833B7 on chromosome 22q12.3-13.2 Contains genes for NCF4 (P40PHOX) protein,cytokine receptor common beta chain precursor CSF2RB (partial), ESTs, CA repeat, STS, GSS /cds=(629,1648) /gb=AL008637 /gi=3136	38893_at
CAT (catalase)	AL035079	Hs.76359	NM_001752	11p13	Cluster Incl. AL035079:dJ53C18.1 (Catalase) /cds=(74,1657) /gb=AL035079 /gi=4775614 /ug=Hs.76359 /len=2287	37009_at
	L43366				L43366 /FEATURE=mRNA /DEFINITION=HUMCADF Homo sapiens (clone jj1b) cadherin mRNA fragment	637_at
TKT (transketolase (Wernicke-Korsakoff syndrome))	L12711	Hs.89643	NM_001064	3p14.3	Cluster Incl. L12711:Homo sapiens transketolase (tk) mRNA, complete cds	38789_at

syndrome))					/cds=(98,1969) /gb=L12711 /gi=388890 /ug=Hs.89643 /len=2069	
STAT6 (signal transducer and activator of transcription 6, interleukin-4 induced)	U16031	Hs.181015	NM_003153	12q13	U16031 /FEATURE= /DEFINITION=HSU16031 Human transcription factor IL-4 Stat mRNA, complete cds	845_at
PLOD (procollagen-lysine, 2-oxoglutarate 5-dioxygenase (lysine hydroxylase, Ehlers-Danlos syndrome type VI))	L06419	Hs.75093	NM_000302	1p36.3-p36.2	Cluster Incl. L06419:Homo sapiens lysyl hydroxylase (PLOD) mRNA, complete cds /cds=(200,2383) /gb=L06419 /gi=190073 /ug=Hs.75093 /len=3115	36184_at
LRP3 (low density lipoprotein receptor-related protein 3)	AB009462	Hs.143641	NM_002333	19q13.1	Cluster Incl. AB009462:Homo sapiens hLRp105 mRNA for LDL receptor related protein 105, complete cds /cds=(226,2538) /gb=AB009462 /gi=3413957 /ug=Hs.143641 /len=2601	31815_r_at
BLCAP (bladder cancer associated protein)	AL049288	Hs.5300	NM_006698	20q11.2-q12	Cluster Incl. AL049288:Homo sapiens mRNA; cDNA DKFZp564M053 (from clone DKFZp564M053) /cds=UNKNOWN	35267_g_at

					/gb=AL049288 /gi=4500049 /ug=Hs.5300 /len=2018	
WBSCR1 (Williams-Beuren syndrome chromosome region 1)	D26068	Hs.180900	NM_022170	7q11.23	Cluster Incl. D26068:Human mRNA for KIAA0038 gene, partial cds /cds=(0,694) /gb=D26068 /gi=436225 /ug=Hs.180900 /len=2477	41212_r_at
PPP6C (protein phosphatase 6, catalytic subunit)	X92972	Hs.80324	NM_002721	xq22.3	Cluster Incl. X92972:H.sapiens mRNA for protein phosphatase 6 /cds=(21,938) /gb=X92972 /gi=5701862 /ug=Hs.80324 /len=1292	37581_at
UNRIP(unr-interacting protein)	AB024327	Hs.3727	NM_007178	12	Cluster Incl. AB024327:Homo sapiens pt-wd mRNA for WD-40 repeat protein, complete cds /cds=(300,1352) /gb=AB024327 /gi=4519416 /ug=Hs.3727 /len=1850	34402_at
EP300 (E1A binding protein p300)	U01877	Hs.25272	NM_001429	22q13.2	Cluster Incl. U01877:Human p300 protein mRNA, complete cds /cds=(1199,8443) /gb=U01877 /gi=495300 /ug=Hs.25272	33896_at

					/len=9046	
DKFZP434D1335(DKFZP434D1335 protein)	AI920820	Hs.8258		19	Cluster Incl. AI920820:wn82e10.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-2452362 /clone_end=3 /gb=AI920820 /gi=5656784 /ug=Hs.8258 /len=519	38400_at
GCGR (glucagon receptor)	L20316	Hs.208	NM_000160	17q25	Cluster Incl. L20316:Human glucagon receptor mRNA, complete cds /cds=(277,1710) /gb=L20316 /gi=405189 /ug=Hs.208 /len=2034	32886_at
KIAA0842(KIAA0842 protein)	AB020649	Hs.74569		1	Cluster Incl. AB020649:Homo sapiens mRNA for KIAA0842 protein, partial cds /cds=(0,3062) /gb=AB020649 /gi=4240172 /ug=Hs.74569 /len=3896	36150_at
USP7 (ubiquitin specific protease 7 (herpes virus-associated))	Z72499	Hs.78683	NM_003470	16p13.3	Cluster Incl. Z72499:H.sapiens mRNA for herpesvirus associated ubiquitin-specific protease (HAUSP) /cds=(199,3507) /gb=Z72499 /gi=1545951 /ug=Hs.78683	37672_at

					/len=4022	
UGTREL7(UDP-glucuronic acid/UDP-N-acetylgalactosamine dual transporter)	D87449	Hs.82635	NM_015139	1	Cluster Incl. D87449:Human mRNA for KIAA0260 gene, partial cds /cds=(0,1153) /gb=D87449 /gi=1665786 /ug=Hs.82635 /len=5918	37888_at
MSF (MLL septin-like fusion (NOTE: non-standard symbol and name))	AB023208	Hs.181002	NM_006640	17q25	Cluster Incl. AB023208:Homo sapiens mRNA for KIAA0991 protein, complete cds /cds=(732,2000) /gb=AB023208 /gi=4589625 /ug=Hs.181002 /len=3938	41220_at
PPP1R8 (protein phosphatase 1, regulatory (inhibitor) subunit 8)	U14575	Hs.78961	NM_002713	1p35	Cluster Incl. U14575:Human (ard-1) mRNA, complete cds /cds=(935,1318) /gb=U14575 /gi=559771 /ug=Hs.78961 /len=2401	37705_at
LNK(linker of T-cell receptor pathways)	AF055581	Hs.13131	NM_005475	12	Cluster Incl. AF055581:Homo sapiens adaptor protein Lnk mRNA, complete cds /cds=(357,2084) /gb=AF055581	39428_at

					/gi=3845720 /ug=Hs.13131 /len=5403	
DYRK1A (dual-specificity tyrosine-(Y)-phosphorylation regulated kinase 1A)	D86550	Hs.75842	NM_001396	21q22.13	D86550 /FEATURE= /DEFINITION=D86550 Human mRNA for serine/threonine protein kinase, complete cds	1512_at
ZNF24 (zinc finger protein 24 (KOX 17))	AF016052	Hs.183593	NM_006965	18q12	Cluster Incl. AF016052:Homo sapiens zinc finger protein ZNF191 (ZNF191) gene, complete cds /cds=(165,1271) /gb=AF016052 /gi=2394173 /ug=Hs.183593 /len=2976	33306_at
FLJ11126(hypothetical protein FLJ11126)	AA034074	Hs.226396	NM_018332	16	Cluster Incl. AA034074:zi06c05.r1 Homo sapiens cDNA, 5 end /clone=IMAGE-429992 /clone_end=5 /gb=AA034074 /gi=1505901 /ug=Hs.226396 /len=655	33394_at
EIF2S3 (eukaryotic translation initiation factor 2, subunit 3 (gamma, 52kD))	L19161	Hs.211539	NM_001415	xp22.2-p22.1	Cluster Incl. L19161:Human translation initiation factor eIF-2 gamma subunit mRNA, complete cds /cds=(0,1418) /gb=L19161 /gi=306899 /ug=Hs.211539	35934_at

					/len=1440	
UBE2N (ubiquitin-conjugating enzyme E2N (homologous to yeast UBC13))	D83004	Hs.75355	NM_003348	12	D83004 /FEATURE= /DEFINITION=D83004 Human epidermoid carcinoma mRNA for ubiquitin-conjugating enzyme E2 similar to Drosophila bendless gene product, complete cds	1660_at
RNF6 (ring finger protein (C3H2C3 type) 6)	AJ010346	Hs.32597	NM_005977	13q12.2	Cluster Incl. AJ010346:Homo sapiens mRNA for RING-H2 protein RNF6, alternative exon 1a /cds=(360,2417) /gb=AJ010346 /gi=4583651 /ug=Hs.32597 /len=3503	35656_at
KIAA0138(KIAA0138 gene product)	D50928	Hs.159384	NM_014649	19	Cluster Incl. D50928:Human mRNA for KIAA0138 gene, complete cds /cds=(36,2897) /gb=D50928 /gi=1469198 /ug=Hs.159384 /len=3233	32099_at
FGFR1 (fibroblast growth factor receptor 1 (fms-related tyrosine kinase 2, Pfeiffer syndrome))	X66945	Hs.748	NM_000604	8p11.2-p11.1	X66945 /FEATURE=cds /DEFINITION=HSNSAMTK H.sapiens N-	424_s_at

related tyrosine kinase 2, Pfeiffer syndrome))					sam mRNA for fibroblast growth factor receptor	
C6orf5 (chromosome 6 open reading frame 5)	AL050289	Hs.7446	NM_015524	6q21	Cluster Incl. AL050289:Homo sapiens mRNA; cDNA DKFZp586G0522 (from clone DKFZp586G0522) /cds=(179,1876) /gb=AL050289 /gi=4886510 /ug=Hs.7446 /len=2364	36139_at
EP300 (E1A binding protein p300)	U01877	Hs.25272	NM_001429	22q13.2	U01877 /FEATURE= /DEFINITION=HSU01877 Human p300 protein mRNA, complete cds	551_at
BRAP (BRCA1 associated protein)	AL042733	Hs.122764	NM_006768	12q24	Cluster Incl. AL042733:DKFZp434B2222_s1 Homo sapiens cDNA, 3 end /clone=DKFZp434B2222 /clone_end=3 /gb=AL042733 /gi=5422182 /ug=Hs.30982 /len=782	41512_at
TGFBR2 (transforming growth factor, beta receptor II (70-80kD))	D50683	Hs.82028	NM_003242	3p22	D50683 /FEATURE= /DEFINITION=D50683 Homo sapiens	1814_at

receptor II (70-80kD))					mRNA for TGF-betaIIR alpha, complete cds	
KIAA0553(KIAA0553 protein)	AB011125	Hs.105749		17	Cluster Incl. AB011125:Homo sapiens mRNA for KIAA0553 protein, partial cds /cds=(0,3289) /gb=AB011125 /gi=3043629 /ug=Hs.105749 /len=5574	38668_at
TGFBR2 (transforming growth factor, beta receptor II (70-80kD))	D50683	Hs.82028	NM_003242	3p22	D50683 /FEATURE= /DEFINITION=D50683 Homo sapiens mRNA for TGF-betaIIR alpha, complete cds	1815_g_at
SFPQ (splicing factor proline/glutamine rich (polypyrimidine tract-binding protein-associated))	X70944	Hs.180610	NM_005066	1pter-p32.3	Cluster Incl. X70944:H.sapiens mRNA for PTB-associated splicing factor /cds=(85,2208) /gb=X70944 /gi=38457 /ug=Hs.180610 /len=3071	40638_at
CHRAC17(DNA polymerase epsilon, subunit 3)	AF070640	Hs.108112	NM_017443	9	Cluster Incl. AF070640:Homo sapiens clone 24781 mRNA sequence /cds=UNKNOWN /gb=AF070640	38702_at

					/gi=3283913 /ug=Hs.108112 /len=1583	
SF3B4 (splicing factor 3b, subunit 4, 49kD)	L35013	Hs.25797	NM_005850	1q12-q21	Cluster Incl. L35013:Human spliceosomal protein (SAP 49) gene, complete cds /cds=(0,1274) /gb=L35013 /gi=556216 /ug=Hs.25797 /len=1275	33909_at
COL6A2 (collagen, type VI, alpha 2)	M20777	Hs.159263		21q22.3	Cluster Incl. M20777:Homo sapiens, alpha-2 (VI) collagen /cds=UNKNOWN /gb=M20777 /gi=180910 /ug=Hs.159263 /len=1005	32098_at
GNA15 (guanine nucleotide binding protein (G protein), alpha 15 (Gq class))	M63904	Hs.73797	NM_002068	19p13.3	Cluster Incl. M63904:Human G-alpha 16 protein mRNA, complete cds /cds=(219,1343) /gb=M63904 /gi=182891 /ug=Hs.73797 /len=2060	40365_at
KIAA0602(KIAA0602 protein)	AB011174	Hs.37656		14	Cluster Incl. AB011174:Homo sapiens mRNA for KIAA0602 protein, partial cds /cds=(0,2889) /gb=AB011174 /gi=3043727 /ug=Hs.37656 /len=3428	34406_at

KIAA0997(KIAA0997 protein)	AI970189	Hs.24083	NM_014950	14	Cluster Incl. AI970189:wr08d01.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-2480929 /clone_end=3 /gb=AI970189 /gi=5767015 /ug=Hs.24083 /len=659	34751_at
TAF2I (TATA box binding protein (TBP)-associated factor, RNA polymerase II, I, 28kD)	X83928	Hs.83126	NM_005643	6	X83928 /FEATURE=cds /DEFINITION=HSTAFII28 H.sapiens mRNA for transcription factor TFIID subunit TAFII28	131_at
MGC4175(hypothetical protein MGC4175)	AI656421	Hs.322404	NM_024315	7	Cluster Incl. AI656421:tt50h10.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-2244259 /clone_end=3 /gb=AI656421 /gi=4740400 /ug=Hs.5671 /len=566	41809_at
RY1(putative nucleic acid binding protein RY-1)	X76302	Hs.54649		2	Cluster Incl. X76302:H.sapiens RY-1 mRNA for putative nucleic acid binding protein /cds=(0,493) /gb=X76302 /gi=431952 /ug=Hs.54649 /len=1402	35286_r_at
CDKN1A (cyclin-dependent kinase inhibitor 1A (p21, Cip1))	U03106	Hs.179665	NM_000389	6p21.2	U03106 /FEATURE= /DEFINITION=HSU03106 Human wild-	2031_s_at

(p21, Cip1))					type p53 activated fragment-1 (WAF1) mRNA, complete cds	
WEE1 (wee1+ (S. pombe) homolog)	W28575	Hs.75188	NM_003390	11p15.3-p15.1	Cluster Incl. W28575:51f12 Homo sapiens cDNA /gb=W28575 /gi=1308730 /ug=Hs.8151 /len=906	38102_at
KIAA0143(KIAA0143 protein)	D63477	Hs.84087		8	Cluster Incl. D63477:Human mRNA for KIAA0143 gene, partial cds /cds=(0,2658) /gb=D63477 /gi=1469867 /ug=Hs.84087 /len=5286	38472_at
SLBP (stem-loop (histone) binding protein)	U75679	Hs.75257	NM_006527	4p16.3	Cluster Incl. U75679:Human histone stem-loop binding protein (SLBP) mRNA, complete cds /cds=(115,927) /gb=U75679 /gi=1732076 /ug=Hs.75257 /len=1725	36913_at
KRN1 (keratin, cuticle, ultrahigh sulphur 1)	X63755	Hs.2743	NM_005553	11q13.5	Cluster Incl. X63755:H.sapiens mRNA for high-sulphur keratin /cds=(238,747) /gb=X63755 /gi=32471 /ug=Hs.2743 /len=1024	34555_at

COPB (coatomer protein complex, subunit beta)	X82103	Hs.3059	NM_016451	11pter-p15.5	Cluster Incl. X82103:H.sapiens mRNA for beta-COP /cds=(0,911) /gb=X82103 /gi=620109 /ug=Hs.3059 /len=1183	34326_at
TAF2I (TATA box binding protein (TBP)-associated factor, RNA polymerase II, I, 28kD)	X83928	Hs.83126	NM_005643	6	Cluster Incl. X83928:H.sapiens mRNA for transcription factor TFIID subunit TAFII28 /cds=(92,727) /gb=X83928 /gi=791056 /ug=Hs.83126 /len=925	38426_at
ZFX (zinc finger protein, X-linked)	X59739	Hs.2074	NM_003410	xp21.3	Cluster Incl. X59739:Human ZFX mRNA for put. transcription activator, isoform 2 /cds=(78,2492) /gb=X59739 /gi=38021 /ug=Hs.2074 /len=5527	38931_at
GOLPH1 (golgi phosphoprotein 1)	AF020762	Hs.6831	NM_022735	1q41	Cluster Incl. AF020762:Homo sapiens clone 1400 unknown protein mRNA, partial cds /cds=(0,805) /gb=AF020762 /gi=2738926 /ug=Hs.6831 /len=1319	36827_at
PTP4A2 (protein tyrosine phosphatase type IVA, member 2)	U14603	Hs.82911	NM_003479	1p35	Cluster Incl. U14603:Human protein-tyrosine phosphatase (HU-PP-1) mRNA, partial sequence /cds=(423,926)	38415_at

					/gb=U14603 /gi=894158 /ug=Hs.82911 /len=1526	
KIAA0863(KIAA0863 protein)	AB020670	Hs.131915	NM_014913	18	Cluster Incl: AB020670:Homo sapiens mRNA for KIAA0863 protein, complete cds /cds=(185,3580) /gb=AB020670 /gi=4240214 /ug=Hs.131915 /len=4313	37837_at
UBE2D3 (ubiquitin-conjugating enzyme E2D 3 (homologous to yeast UBC4/5))	U39318	Hs.118797	NM_003340	4q24-q26	Cluster Incl. U39318:Human E2 ubiquitin conjugating enzyme UbcH5C (UBCH5C) mRNA, complete cds /cds=(45,488) /gb=U39318 /gi=1145690 /ug=Hs.118797 /len=724	39083_at
NTPBP(XPA binding protein 1; putative ATP(GTP)-binding protein)	AJ010842	Hs.18259	NM_007266	2	Cluster Incl. AJ010842:Homo sapiens mRNA for putative ATP(GTP)-binding protein, partial /cds=(0,1077) /gb=AJ010842 /gi=3646129 /ug=Hs.18259 /len=1722	41756_at

UBE2N (ubiquitin-conjugating enzyme E2N (homologous to yeast UBC13))	D83004	Hs.75355	NM_003348	12	Cluster Incl. D83004:Human epidermoid carcinoma mRNA for ubiquitin-conjugating enzyme E2 similar to Drosophila bendless gene product, complete cds /cds=(63,521) /gb=D83004 /gi=1181557 /ug=Hs.75355 /len=1203	36604_at
NCOR2 (nuclear receptor co-repressor 2)	U37146	Hs.287994	NM_006312	12q24	Cluster Incl. U37146:Human silencing mediator of retinoid and thyroid hormone action (SMRT) mRNA, complete cds /cds=(495,4982) /gb=U37146 /gi=1045654 /ug=Hs.120980 /len=5970	39358_at
LGALS2 (lectin, galactoside-binding, soluble, 2 (galectin 2) (NOTE: redefinition of symbol))	AL022315	Hs.113987	NM_006498	22q13.1	Cluster Incl. AL022315:dJ1177I5.3 (Lectin, Galactose-binding, soluble, 2 (Galectin 2, S-Lac Lectin 2, HL14)) /cds=(80,478) /gb=AL022315 /gi=3820991 /ug=Hs.113987 /len=494	37456_at
ERPROT213-21(protein with polyglutamine repeat; calcium (ca2+) homeostasis endoplasmic reticulum protein)	U94836	Hs.6430	NM_006387	19	Cluster Incl. U94836:Human ERPROT 213-21 mRNA, complete cds /cds=(88,2742) /gb=U94836 /gi=2058690	41836_at

endoplasmic reticulum protein)					/ug=Hs.6430 /len=4003	
FMR1 (fragile X mental retardation 1)	X69962	Hs.89764	NM_002024	xq27.3	Cluster Incl. X69962:H.sapiens FMR-1 mRNA /cds=(219,2117) /gb=X69962 /gi=296587 /ug=Hs.89764 /len=4362	37994_at
SUPV3L1 (suppressor of var1 (S.cerevisiae) 3-like 1)	AF042169	Hs.106469	NM_003171	10q22.1	Cluster Incl. AF042169:Homo sapiens putative ATP-dependent mitochondrial RNA helicase (SUV3) mRNA, nuclear gene encoding mitochondrial protein, complete cds /cds=(0,2360) /gb=AF042169 /gi=2801554 /ug=Hs.106469 /len=2361	41408_at
NUFIP1 (nuclear fragile X mental retardation protein interacting protein 1)	AL049842	Hs.120247	NM_012345	13q14	Cluster Incl. AL049842:Human DNA sequence from clone 129L7 on chromosome 6q12-13. Contains the gene for a PUTATIVE novel protein, ESTs, an STS, GSSs and a taga repeat polymorphism /cds=(9,749) /gb=AL049842 /gi=5419768 /ug=Hs.120247 /len=1679	37518_at

GNE(UDP-N-acetylglucosamine-2-epimerase/N-acetylmannosamine kinase)	AJ238764	Hs.5920	NM_005476	9	Cluster Incl. AJ238764:Homo sapiens mRNA for UDP-N-acetylglucosamine-2-epimerase_/ N-acetylmannosamine kinase /cds=(41,2209) /gb=AJ238764 /gi=4775361 /ug=Hs.5920 /len=3649	36515_at
SAS10(disrupter of silencing 10)	AI126004	Hs.322901	NM_020368	4	Cluster Incl. AI126004:qc50e12.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-1713070 /clone_end=3 /gb=AI126004 /gi=3594518 /ug=Hs.87627 /len=611	33150_at
UBE3A (ubiquitin protein ligase E3A (human papilloma virus E6-associated protein, Angelman syndrome))	U84404	Hs.180686	NM_000462	15q11-q13	Cluster Incl. U84404:Human E6-associated protein E6-AP/ubiquitin-protein ligase (UBE3A) mRNA, alternatively spliced, complete cds /cds=(586,3144) /gb=U84404 /gi=1872513 /ug=Hs.180686 /len=3168	41205_at
TRAM(translocating chain-associating membrane protein)	X63679	Hs.4147	NM_014294	8	Cluster Incl. X63679:H.sapiens mRNA for TRAMP protein /cds=(121,1245) /gb=X63679 /gi=37264 /ug=Hs.4147	34796_at

					/len=1267	
PRCC (papillary renal cell carcinoma (translocation-associated))	X99720	Hs.9629	NM_005973	1q21.1	Cluster Incl. X99720:H.sapiens TPRC gene /cds=(212,1687) /gb=X99720 /gi=1869817 /ug=Hs.9629 /len=2053	39149_at
HBOA(histone acetyltransferase)	AI951946	Hs.21907	NM_007067	X	Cluster Incl. AI951946:wx39f10.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-2546059 /clone_end=3 /gb=AI951946 /gi=5744256 /ug=Hs.244 /len=523	41338_at
SNRPA1 (small nuclear ribonucleoprotein polypeptide A')	X13482	Hs.80506	NM_003090	22q	Cluster Incl. X13482:Human mRNA for U2 snRNP-specific A protein /cds=(56,823) /gb=X13482 /gi=37546 /ug=Hs.80506 /len=1033	37585_at
ZFR(zinc finger RNA binding protein)	AI743507	Hs.173518	NM_016107	5	Cluster Incl. AI743507:wf72a06.x2 Homo sapiens cDNA, 3 end /clone=IMAGE-2361106 /clone_end=3 /gb=AI743507 /gi=5111795 /ug=Hs.173518 /len=733	40610_at

RAGA(Ras-related GTP-binding protein)	U41654	Hs.57304	NM_006570	9	Cluster Incl. U41654:Human adenovirus protein E3-14.7k interacting protein 1 (FIP-1) mRNA, complete cds /cds=(243,1184) /gb=U41654 /gi=2058395 /ug=Hs.57304 /len=1610	35316_at
RAC1 (ras-related C3 botulinum toxin substrate 1 (rho family, small GTP binding protein Rac1))	D25274	Hs.173737	NM_006908	7p22	Cluster Incl. D25274:Homo sapiens mRNA, clone-PO2ST9 /cds=UNKNOWN /gb=D25274 /gi=464185 /ug=Hs.173737 /len=1232	40864_at
	AB015202				Cluster Incl. AB015202:Homo sapiens gene for hippocalcin /cds=(235,816) /gb=AB015202 /gi=4417205 /ug=Hs.114215 /len=1584	41602_at
PFDN4 (prefoldin 4)	U41816	Hs.91161	NM_002623	20q13	Cluster Incl. U41816:Human C-1 mRNA, complete cds /cds=(11,403) /gb=U41816 /gi=1620560 /ug=Hs.91161 /len=1203	41003_at
YY1 (YY1 transcription factor)	M77698	Hs.97496	NM_003403	14q	M77698 /FEATURE= /DEFINITION=HUMKRP Homo sapiens	891_at

					GLI-Krupple related protein (YY1) mRNA, complete cds	
REM(GTPase GES; REM protein)	AF084465	Hs.247729	NM_014012	20	Cluster Incl. AF084465:Homo sapiens Ras-like GTP-binding protein REM mRNA, complete cds /cds=(72,968) /gb=AF084465 /gi=3462895 /ug=Hs.87062 /len=976	34008_at
CLTC (clathrin, heavy polypeptide (Hc))	D21260	Hs.178710	NM_004859	17q11-qter	Cluster Incl. D21260:Human mRNA for KIAA0034 gene, complete cds /cds=(172,5199) /gb=D21260 /gi=434760 /ug=Hs.178710 /len=6111	41159_at
	AL049229				Cluster Incl. AL049229:Homo sapiens mRNA; cDNA DKFZp564O1016 (from clone DKFZp564O1016) /cds=UNKNOWN /gb=AL049229 /gi=4499961 /ug=Hs.15787 /len=1767	32082_at
BAG5 (BCL2-associated athanogene 5)	AB020680	Hs.5443	NM_004873	14	Cluster Incl. AB020680:Homo sapiens mRNA for KIAA0873 protein, partial cds	36463_at

					/cds=(0,1400) /gb=AB020680 /gi=4240234 /ug=Hs.5443 /len=4119	
HSPA9B (heat shock 70kD protein 9B (mortalin-2))	L15189	Hs.3069	NM_004134	5q31.1	Cluster Incl. L15189:Homo sapiens mitochondrial HSP75 mRNA, complete cds /cds=(29,2068) /gb=L15189 /gi=292058 /ug=Hs.3069 /len=2131	41510_s_at
PPP2CA (protein phosphatase 2 (formerly 2A), catalytic subunit, alpha isoform)	M60483	Hs.91773	NM_002715	5q23-q31	M60483 /FEATURE=mRNA /DEFINITION=HUMPP2AA Human protein phosphatase 2A catalytic subunit-alpha gene, complete cds	237_s_at
TERF1 (telomeric repeat binding factor (NIMA-interacting) 1)	U74382	Hs.194562	NM_003218	8q13	U74382 /FEATURE= /DEFINITION=HSU74382 Human telomeric repeat DNA-binding protein (PIN2) mRNA, complete cds	1329_s_at
KIAA0685(KIAA0685 gene product)	AI677689	Hs.296406	NM_014678	22	Cluster Incl. AI677689:wd33c06.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-2329930 /clone_end=3 /gb=AI677689	40222_s_at

					/gi=4887871 /ug=Hs.153121 /len=478	

Table 11:

UCL/HGNC/HUGO Human Gene Nomenclature Database Symbol	GenBank Accession No.	UniGene Cluster	RefSeq	Chromosomal Location	Description Unigene Build #95	Gene Name
NCOA1 (nuclear receptor coactivator 1	AJ000882	Hs.74002	NM_003743	2p23	Cluster Incl. AJ000882:Homo sapiens mRNA for steroid receptor coactivator 1e /cds=(201,4400) /gb=AJ000882 /gi=2924310 /ug=Hs.74002 /len=4709	36118_at
NCOA1 (nuclear receptor coactivator 1	U59302	Hs.74002	NM_003743	2p23	U59302 /FEATURE= /DEFINITION=HSU59302 Human steroid receptor coactivator-1 F-SRC-1 mRNA, complete cds	484_at
FCGR2B (Fc fragment of IgG, low affinity IIb, receptor for (CD32)	M28696	Hs.278443	NM_004001	1q23	Cluster Incl. M28696:Human low-affinity IgG Fc receptor (beta-Fc-gamma-RII) mRNA, complete cds /cds=(41,916) /gb=M28696 /gi=184843 /ug=Hs.233450	34663_at

					/len=1416	
RBL2 (retinoblastoma-like 2 (p130))	X76061	Hs.79362	NM_005611	16q12.2	Cluster Incl. X76061:H.sapiens p130 mRNA for 130K protein /cds=(69,3488) /gb=X76061 /gi=416030 /ug=Hs.79362 /len=4835	32597_at
	AI749193				Cluster Incl. AI749193:at40e04.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-2374494 /clone_end=3 /gb=AI749193 /gi=5127457 /ug=Hs.17639 /len=544	40623_at
ITGB7 (integrin, beta 7	M68892	Hs.1741	NM_000889	12q13.13	M68892 /FEATURE= /DEFINITION=HUMINTB7 Human integrin beta-7 subunit mRNA, complete cds	2019_s_at
ITGB7 (integrin, beta 7)	U50534	Hs.1741	NM_000889	12q13.13	U50534 /FEATURE= /DEFINITION=HSU50534 Human BRCA2 region, mRNA sequence CG003	1529_at

KIAA0476(KIAA0476 gene product	AB007945	Hs.6684	NM_014856	1	Cluster Incl. AB007945:Homo sapiens mRNA for KIAA0476 protein, complete cds /cds=(568,4728) /gb=AB007945 /gi=3413913 /ug=Hs.6684 /len=5525	35786_at
KIAA0240(KIAA0240 protein	D87077	Hs.196275		6	Cluster Incl. D87077:Human mRNA for KIAA0240 gene, partial cds /cds=(0,2953) /gb=D87077 /gi=1510154 /ug=Hs.196275 /len=6060	38892_at
UCP2 (uncoupling protein 2 (mitochondrial, proton carrier	U94592	Hs.80658	NM_003355	11q13	Cluster Incl. U94592:Human uncoupling protein homolog (UCPH) mRNA, complete cds /cds=(314,1243) /gb=U94592 /gi=2052354 /ug=Hs.80658 /len=1888	37591_at
ADAM19 (a disintegrin and metalloproteinase domain 19 (meltrin beta))	AL049415	Hs.278679	NM_023038	5q32-q33	Cluster Incl. AL049415:Homo sapiens mRNA; cDNA DKFZp586N2119 (from clone DKFZp586N2119) /cds=UNKNOWN /gb=AL049415 /gi=4500196 /ug=Hs.204290 /len=1232	33812_at

CDW52 (CDW52 antigen (CAMPATH-1 antigen))	N90866	Hs.276770	NM_001803	1p36	Cluster Incl. N90866:zb11b10.s1 Homo sapiens cDNA, 3 end /clone=IMAGE-301723 /clone_end=3 /gb=N90866 /gi=1444193 /ug=Hs.214742 /len=577	34210_at
CD37 (CD37 antigen)	X14046	Hs.153053	NM_001774	19p13-q13.4	Cluster Incl. X14046:Human mRNA for leukocyte antigen CD37 /cds=(63,908) /gb=X14046 /gi=29793 /ug=Hs.153053 /len=1125	31870_at
IRF6 (interferon regulatory factor 6	AL022398	Hs.11801	NM_006147	1q32.3-q41	Cluster Incl. AL022398:dJ434O14.3.1 (putative protein) (isoform 1) /cds=(0,988) /gb=AL022398 /gi=3355547 /ug=Hs.87684 /len=1241	40721_g_at
KIAA1002(KIAA1002 protein	AB023219	Hs.20340			Cluster Incl. AB023219:Homo sapiens mRNA for KIAA1002 protein, complete cds /cds=(800,3322) /gb=AB023219 /gi=4589647 /ug=Hs.102483 /len=4331	41366_at
TGFBR3 (transforming growth factor, beta receptor III (betaglycan, 300kD)	L07594	Hs.79059	NM_003243	1p33-p32	L07594 /FEATURE= /DEFINITION=HUMTGFB3C Human	1897_at

receptor III (betaglycan, 300kD)					transforming growth factor-beta type III receptor (TGF-beta) mRNA, complete cds	
LAMA5 (laminin, alpha 5	AB011105	Hs.11669	NM_005560	20q13.2-q13.3	Cluster Incl. AB011105:Homo sapiens mRNA for KIAA0533 protein, partial cds /cds=(0,4939) /gb=AB011105 /gi=3043589 /ug=Hs.11669 /len=5117	41610_at
FLJ10140(hypothetical protein FLJ10140	AL031588	Hs.250671	NM_018006	22	Cluster Incl. AL031588:dJ1163J1.1 (ortholog of mouse transmembrane receptor Celsr1 (KIAA0279 LIKE EGF-like domain containing protein similar to rat MEG /cds=(0,4433) /gb=AL031588 /gi=4007108 /ug=Hs.123043 /len=6438	41660_at
PRKCB1 (protein kinase C, beta 1	X07109	Hs.77202	NM_002738	16p11.2	X07109 /FEATURE=cds /DEFINITION=HSPKCB2A Human mRNA for protein kinase C (PKC) type beta II /NOTE=replacement of probe set 1216_at	160029_at
PLCE2 (phospholipase C, epsilon 2)	AB029015	Hs.54886		3p25.3-p25.1	Cluster Incl. AB029015:Homo sapiens mRNA for KIAA1092 protein, partial cds	41796_at

					/cds=(0,3464) /gb=AB029015 /gi=5689520 /ug=Hs.54886 /len=4147	
C9ORF10(C9orf10 protein)	D80005	Hs.76666	NM_014612	9	Cluster Incl. D80005:Human mRNA for KIAA0183 gene, partial cds /cds=(0,3190) /gb=D80005 /gi=1136425 /ug=Hs.76666 /len=4905	37031_at
TMEM5 (transmembrane protein 5)	AB015633	Hs.112986		1	Cluster Incl. AB015633:Homo sapiens mRNA for type II membrane protein, complete cds, clone-HP10481 /cds=(104,1435) /gb=AB015633 /gi=4586843 /ug=Hs.112986 /len=1451	37445_at
TUCAN(tumor up-regulated CARD-containing antagonist of caspase nine)	AB023172			19	Cluster Incl. AB023172:Homo sapiens mRNA for KIAA0955 protein, complete cds /cds=(313,1608) /gb=AB023172 /gi=4589553 /ug=Hs.10031 /len=5059	41100_at
PFTK1 (PFTAIRE protein kinase 1	AB020641	Hs.57856	NM_012395	7q21-q22	Cluster Incl. AB020641:Homo sapiens mRNA for KIAA0834 protein, complete cds /cds=(144,1499) /gb=AB020641	36502_at

					/gi=4240156 /ug=Hs.57856 /len=4957	
IRF6 (interferon regulatory factor 6	AL022398	Hs.11801	NM_006147	1q32.3-q41	Cluster Incl. AL022398:dJ434O14.3.3 (novel protein) (isoform 3) /cds=(290,1885) /gb=AL022398 /gi=3355547 /ug=Hs.87684 /len=2058	40719_at
M17S2 (membrane component, chromosome 17, surface marker 2 (ovarian carcinoma antigen	D30756	Hs.277721	NM_005899	17q21.1	Cluster Incl. D30756:Human mRNA for KIAA0049 gene, complete cds /cds=(140,3040) /gb=D30756 /gi=488500 /ug=Hs.233745 /len=4654	33444_at
DKFZP564K0822(hypothetical protein DKFZp564K0822	W25986	Hs.4750	NM_030796	7	Cluster Incl. W25986:17e7 Homo sapiens cDNA /gb=W25986 /gi=1306253 /ug=Hs.4750 /len=769	34830_at
KIAA0430(KIAA0430 gene produc	AB007890		-	16	Cluster Incl. AB007890:Homo sapiens KIAA0430 mRNA, complete cds /cds=(0,3172) /gb=AB007890 /gi=2887438 /ug=Hs.166163 /len=6011	31936_s_at

KIAA1696(KIAA1696 protein)	N98667	Hs.106826	NM_016621	11	Cluster Incl. N98667:yy66d05.r1 Homo sapiens cDNA, 5 end /clone=IMAGE-278505 /clone_end=5 /gb=N98667 /gi=1270089 /ug=Hs.106826 /len=549	39551_at
GABBR1 (gamma-aminobutyric acid (GABA) B receptor, 1	AJ225028	Hs.167017	NM_001470	6p21.3	Cluster Incl. AJ225028:Homo sapiens mRNA for GABA-B R1a receptor /cds=(234,3119) /gb=AJ225028 /gi=3892593 /ug=Hs.167017 /len=4434	32623_at
OGDH (oxoglutarate dehydrogenase (lipoamide)	D10523	Hs.168669	NM_002541	7p14-p13	Cluster Incl. D10523:Human mRNA for 2-oxoglutarate dehydrogenase, complete cds /cds=(57,3065) /gb=D10523 /gi=531240 /ug=Hs.168669 /len=4122	40470_at
CBX7 (chromobox homolog 7	AL031846			22q13.1	Cluster Incl. AL031846:dJ742C19.5 (novel Chromobox protein) /cds=(89,844) /gb=AL031846 /gi=4164368 /ug=Hs.7442 /len=3964	36894_at

SP140(nuclear body protein Sp140	U36500	Hs.309943	NM_007237	2	Cluster Incl. U36500:Human lymphoid-specific SP100 homolog (LYSP100-B) mRNA, complete cds /cds=(116,2764) /gb=U36500 /gi=1173653 /ug=Hs.85283 /len=3252	40700_at
13CDNA73(putative gene product	U50534	Hs.181304	NM_023037	13	U50534 /FEATURE= /DEFINITION=HSU50534 Human BRCA2 region, mRNA sequence CG003	1530_g_at
SIAT1 (sialyltransferase 1 (beta-galactoside alpha-2,6-sialytransferase	X62822	Hs.2554	NM_003032	3q27-q28	Cluster Incl. X62822:H.sapiens gene encoding beta-galactoside alpha-2,6-sialyltransferase /cds=(310,1530) /gb=X62822 /gi=29433 /ug=Hs.2554 /len=2699	41352_at
MAP3K5 (mitogen-activated protein kinase kinase kinase 5)	U67156	Hs.151988	NM_005923	6q22.33	U67156 /FEATURE= /DEFINITION=HSU67156 Human mitogen-activated kinase kinase kinase 5 (MAPKKK5) mRNA, complete cds	1327_s_at

KIAA0747(KIAA0747 protein)	AB018290	Hs.8309	NM_015292	12	Cluster Incl. AB018290:Homo sapiens mRNA for KIAA0747 protein, partial cds /cds=(0,3219) /gb=AB018290 /gi=3882214 /ug=Hs.8309 /len=4026	38424_at
PRKCB1 (protein kinase C, beta 1	X07109	Hs.77202	NM_002738	16p11.2	X07109 /FEATURE=cds /DEFINITION=HSPKCB2A Human mRNA for protein kinase C (PKC) type beta II	1217_g_at
KIAA0274(KIAA0274 gene product	D87464	Hs.10037	NM_014845	6	Cluster Incl. D87464:Human mRNA for KIAA0274 gene, complete cds /cds=(124,2847) /gb=D87464 /gi=1665812 /ug=Hs.10037 /len=3010	41101_at
KIAA0660(ras-GTPase-activating protein (GAP<120>) SH3-domain-binding protein 2	AB014560	Hs.6727	NM_012297	4	Cluster Incl. AB014560:Homo sapiens mRNA for KIAA0660 protein, complete cds /cds=(120,1568) /gb=AB014560 /gi=3327133 /ug=Hs.6727 /len=4210	35793_at
FCER2 (Fc fragment of IgE, low affinity II, receptor for (CD23A)	M15059	Hs.1416	NM_002002	19p13.3	Cluster Incl. M15059:Human Fc-epsilon receptor (IgE receptor) mRNA, complete cds (H107 epitope) /cds=(213,1178)	34960_g_at

					/gb=M15059 /gi=182447 /ug=Hs.1416 /len=1530	
HLA-DMB (major histocompatibility complex, class II, DM beta	U15085	Hs.1162	NM_002118	6p21.3	Cluster Incl. U15085:Human HLA-DMB mRNA, complete cds /cds=(233,1024) /gb=U15085 /gi=557701 /ug=Hs.1162 /len=1362	41609_at
KIAA1093(KIAA1093 protein	AB029016	Hs.117333		22	Cluster Incl. AB029016:Homo sapiens mRNA for KIAA1093 protein, partial cds /cds=(0,3613) /gb=AB029016 /gi=5689522 /ug=Hs.117333 /len=4159	37487_at
IGHM (immunoglobulin heavy constant mu)	X58529	Hs.302063		14q32.33	Cluster Incl. X58529:Human rearranged immunoglobulin mRNA for mu heavy chain enhancer and constant region /cds=UNKNOWN /gb=X58529 /gi=33480 /ug=Hs.179543 /len=2325	41166_at
KIAA0649(KIAA0649 gene product	AB014549	Hs.26163	NM_014811	9	Cluster Incl. AB014549:Homo sapiens mRNA for KIAA0649 protein, complete cds /cds=(549,4178) /gb=AB014549	39580_at

					/gi=3327111 /ug=Hs.26163 /len=4932	
KIAA0494(KIAA0494 gene product	AB007963	Hs.62515	NM_014774	1	Cluster Incl. AB007963:Homo sapiens mRNA for KIAA0494 protein, complete cds /cds=(977,2464) /gb=AB007963 /gi=3413937 /ug=Hs.62515 /len=5766	41830_at
CLOCK (clock (mouse) homolog	AB002332	Hs.50722	NM_004898	4q12	Cluster Incl. AB002332:Human mRNA for KIAA0334 gene, complete cds /cds=(251,2791) /gb=AB002332 /gi=2224608 /ug=Hs.50722 /len=5715	36080_at
JIK(STE20-like kinase)	AA576724	Hs.12040	NM_016281	12	Cluster Incl. AA576724:nm81b04.s1 Homo sapiens cDNA, 3 end /clone=IMAGE-1074607 /clone_end=3 /gb=AA576724 /gi=2354198 /ug=Hs.12040 /len=580	41646_at
SETBP1 (SET binding protein 1)	AB022660	Hs.151717	NM_015559	18q21.1	Cluster Incl. AB022660:Homo sapiens mRNA for SET-binding protein (SEB), complete cds /cds=(5,4633) /gb=AB022660 /gi=5478317	34990_at

					/ug=Hs.151717 /len=5744	
STAT6 (signal transducer and activator of transcription 6, interleukin-4 induced	AF067575	Hs.181015	NM_003153	12q13	Cluster Incl. AF067575:untitled /cds=(21,2564) /gb=AF067575 /gi=3789867 /ug=Hs.181015 /len=3725	41222_at
DOC2B (double C2-like domains, beta)	D70830	Hs.54402	NM_003585	17	Cluster Incl. D70830:Homo sapiens mRNA for Doc2 beta, complete cds /cds=(160,1398) /gb=D70830 /gi=1235721 /ug=Hs.54402 /len=2043	32422_at
	AA868268				Cluster Incl. AA868268:ak40a05.s1 Homo sapiens cDNA, 3 end /clone=IMAGE-1408400 /clone_end=3 /gb=AA868268 /gi=2963713 /ug=Hs.170267 /len=570	40574_at
	AB018272				Cluster Incl. AB018272:Homo sapiens mRNA for KIAA0729 protein, partial cds /cds=(0,3591) /gb=AB018272 /gi=3882178 /ug=Hs.180948 /len=4143	41218_at

TNFRSF7 (tumor necrosis factor receptor superfamily, member 7	M63928	Hs.180841	NM_001242	12p13	Cluster Incl. M63928:Homo sapiens T cell activation antigen (CD27) mRNA, complete cds /cds=(100,882) /gb=M63928 /gi=180084 /ug=Hs.180841 /len=1204	38578_at
PLCG2 (phospholipase C, gamma 2 (phosphatidylinositol-specific	M37238	Hs.75648	NM_002661	16q24.1	M37238 /FEATURE=mRNA /DEFINITION=HUMPLC Human phospholipase C mRNA, complete cds	1085_s_at
	H24861				Cluster Incl. H24861:yl42e11.r1 Homo sapiens cDNA, 5 end /clone=IMAGE-160940 /clone_end=5 /gb=H24861 /gi=893760 /ug=Hs.90145 /len=517	33168_at
KIAA0543(KIAA0543 protein	AB011115	Hs.98507	H12985S1	7	Cluster Incl. AB011115:Homo sapiens mRNA for KIAA0543 protein, partial cds /cds=(0,3336) /gb=AB011115 /gi=3043609 /ug=Hs.98507 /len=6443	41077_at
HSD17B4 (hydroxysteroid (17-beta) dehydrogenase 4)	X87176	Hs.75441	NM_000414	5q21	Cluster Incl. X87176:H.sapiens mRNA for 17-beta-hydroxysteroid dehydrogenase /cds=(48,2258) /gb=X87176 /gi=1050516	36626_at

					/ug=Hs.75441 /len=2593	
TLK1 (tousled-like kinase 1)	D50927	Hs.18895	NM_012290	8p22-p12	Cluster Incl. D50927:Human mRNA for KIAA0137 gene, complete cds /cds=(1088,2737) /gb=D50927 /gi=1469196 /ug=Hs.18895 /len=4454	32219_at
CASK (calcium/calmodulin-dependent serine protein kinase (MAGUK family)	AF035582	Hs.151469	NM_003688	xp11.4	Cluster Incl. AF035582:Homo sapiens CASK mRNA, complete cds /cds=(15,2708) /gb=AF035582 /gi=2661105 /ug=Hs.151469 /len=3919	31854_at
TRIAD3(TRIAD3 protein	AA650210	Hs.66228	NM_019011	7	Cluster Incl. AA650210:ns88b12.s1 Homo sapiens cDNA /clone=IMAGE-1190687 /gb=AA650210 /gi=2577538 /ug=Hs.116406 /len=528	37476_at
SCAP1 (src family associated phosphoprotein 1	Y11215	Hs.19126	NM_003726	17q21.3	Cluster Incl. Y11215:Homo sapiens mRNA for SKAP55 protein /cds=(70,1149) /gb=Y11215 /gi=2252495 /ug=Hs.19126 /len=1524	38862_at

KIAA0746(KIAA0746 protein	AB018289	Hs.49500		4	Cluster Incl. AB018289:Homo sapiens mRNA for KIAA0746 protein, partial cds /cds=(0,3091) /gb=AB018289 /gi=3882212 /ug=Hs.49500 /len=4086	41585_at
DKFZP586F2423(hypothetical protein DKFZp586F2423	AL080209	Hs.13659		7	Cluster Incl. AL080209:Homo sapiens mRNA; cDNA DKFZp586F2423 (from clone DKFZp586F2423) /cds=UNKNOWN /gb=AL080209 /gi=5262698 /ug=Hs.13659 /len=4241	39692_at
DKFZP434C171(DKFZP434C171 protein	AL080169	Hs.209100	NM_015621	5	Cluster Incl. AL080169:Homo sapiens mRNA; cDNA DKFZp434C171 (from clone DKFZp434C171) /cds=(0,544) /gb=AL080169 /gi=5262637 /ug=Hs.209100 /len=2595	34183_at
ATRX (alpha thalassemia/mental retardation syndrome X-linked (RAD54 (S. cerevisiae) homolog)	U72936	Hs.96264	NM_000489	xq13.1-q21.1	Cluster Incl. U72936:Human putative DNA dependent ATPase and helicase (ATRX) mRNA, alternatively spliced product 1, complete cds /cds=(945,7811) /gb=U72936 /gi=1778306 /ug=Hs.96264	39147_g_at

					/len=10448	
MAPK3 (mitogen-activated protein kinase 3)	X60188	Hs.861		16p12-p11.2	X60188 /FEATURE=mRNA /DEFINITION=HSERK1 Human ERK1 mRNA for protein serine/threonine kinase	1000_at
IGHM (immunoglobulin heavy constant mu)	X67301	Hs.302063		14q32.33	Cluster Incl. X67301:H.sapiens mRNA for IgM heavy chain constant region (Ab63) /cds=(0,1361) /gb=X67301 /gi=38407 /ug=Hs.179543 /len=1453	41165_g_at
	W30677				Cluster Incl. W30677:zb75h10.r1 Homo sapiens cDNA, 5 end /clone=IMAGE-309475 /clone_end=5 /gb=W30677 /gi=1311730 /ug=Hs.5019 /len=614	34871_at
AP1G2 (adaptor-related protein complex 1, gamma 2 subunit)	AI741833	Hs.8991	NM_003917	14q11.2-14q21.3	Cluster Incl. AI741833:wg29e04.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-2366526 /clone_end=3 /gb=AI741833 /gi=5110121 /ug=Hs.8991 /len=658	38798_s_at

CDC4L (cell division cycle 4-like)	M83822	Hs.62354		4	Cluster Incl. M83822:Human beige-like protein (BGL) mRNA, partial cds /cds=(0,5758) /gb=M83822 /gi=1580780 /ug=Hs.62354 /len=7332	35371_at
DKFZP564B116(DKFZP564B116 protein	AL050018	Hs.7387		6	Cluster Incl. AL050018:Homo sapiens mRNA; cDNA DKFZp564B116 (from clone DKFZp564B116) /cds=(0,1151) /gb=AL050018 /gi=4884085 /ug=Hs.7387 /len=2335	36875_at
CSK (c-src tyrosine kinase)	X59932	Hs.77793	NM_004383	15q23-q25	X59932 /FEATURE=mRNA /DEFINITION=HSCSRCKIN Human mRNA for C-SRC-kinase	1768_s_at
LOC54103(hypothetical protein	AL079277	Hs.12969		7	Cluster Incl. AL079277:Homo sapiens mRNA full length insert cDNA clone EUROIMAGE 293605 /cds=(0,806) /gb=AL079277 /gi=5102581 /ug=Hs.12969 /len=1414	41710_at

DKFZP586A0522(DKFZP586A0522 protein	AL050159	Hs.288771	NM_014033	12	Cluster Incl. AL050159:Homo sapiens mRNA; cDNA DKFZp586A0522 (from clone DKFZp586A0522) /cds=(0,732) /gb=AL050159 /gi=4884371 /ug=Hs.108740 /len=1846	38717_at
DGKA (diacylglycerol kinase, alpha (80kD)	X62535	Hs.172690	NM_001345	12q13.3	Cluster Incl. X62535:H.sapiens mRNA for diacylglycerol kinase /cds=(103,2310) /gb=X62535 /gi=30822 /ug=Hs.172690 /len=2564	32716_at
	AL049701				Cluster Incl. AL049701:Human gene from PAC 433G19, chromosome 1 /cds=(0,370) /gb=AL049701 /gi=4678835 /ug=Hs.107325 /len=648	34446_at
IFI41 (interferon-induced protein 41, 30kD)	L22342	Hs.241510	NM_004509		Cluster Incl. L22342:Human nuclear phosphoprotein mRNA, complete cds /cds=(0,746) /gb=L22342 /gi=402204 /ug=Hs.38125 /len=835	35718_at

	AF038199				Cluster Incl. AF038199:Homo sapiens clone 23728 mRNA sequence /cds=UNKNOWN /gb=AF038199 /gi=2795920 /ug=Hs.153106 /len=1112	38154_at
HLA-DOB (major histocompatibility complex, class II, DO beta	X03066	Hs.1802	NM_002120	6p21.3	Cluster Incl. X03066:Human mRNA for HLA-D class II antigen DO beta chain /cds=(56,877) /gb=X03066 /gi=32271 /ug=Hs.1802 /len=1322	38570_at
GTF2E2 (general transcription factor IIE, polypeptide 2 (beta subunit, 34kD)	X63469	Hs.77100	NM_002095	8p21-p12	Cluster Incl. X63469:H.sapiens mRNA for transcription factor TFIIE beta /cds=(242,1117) /gb=X63469 /gi=37069 /ug=Hs.77100 /len=1515	37295_at
NIFU(nitrogen fixation cluster-like	U47101	Hs.9908		12	Cluster Incl. U47101:Human NifU-like protein (hNifU) mRNA, partial cds /cds=(0,366) /gb=U47101 /gi=1685101 /ug=Hs.9908 /len=819	39165_at

LY117 (lymphocyte antigen 117)	AF031137	Hs.88411	NM_007161	6p21.3	Cluster Incl. AF031137:Homo sapiens 1C7 precursor, mRNA, alternatively spliced, complete cds /cds=(264,869) /gb=AF031137 /gi=2623874 /ug=Hs.88411 /len=1041	37968_at
ENTPD6 (ectonucleoside triphosphate diphosphohydrolase 6 (putative function)	AL035252	Hs.12330	NM_001247	20q11.2	Cluster Incl. AL035252:Human DNA sequence from clone 738P15 on chromosome 20p11.2-11.22. Contains a putative new gene, the CD39L2 for nucleoside phosphatase D39-like 2, and the (putative?) IL-6SAG gene in the CD39L2 3 UTR. Contains ESTs, an STS, GSSs and a putative CpG island /cds=(147,1601) /gb=AL035252 /gi=4490906 /ug=Hs.12330 /len=2729	39876_at
ATRX (alpha thalassemia/mental retardation syndrome X-linked (RAD54 (S. cerevisiae) homolog	U72936	Hs.96264	NM_000489	xq13.1-q21.1	U72936 /FEATURE= /DEFINITION=HSU72936 Homo sapiens putative DNA dependent ATPase and helicase (ATRX) mRNA, alternatively spliced product 1, complete cds	818_s_at

MAN2A1 (mannosidase, alpha, class 2A, member 1	D63998	Hs.32965	NM_002372	5q21-q22	Cluster Incl. D63998:Human mRNA for golgi alpha-mannosidaseII, complete cds /cds=(510,3941) /gb=D63998 /gi=974733 /ug=Hs.32965 /len=4101	39663_at
APOC4 (apolipoprotein C-IV)	U32576	Hs.110675	NM_001646	19q13.2	Cluster Incl. U32576:Human apolipoprotein apoC-IV (APOC4) gene, complete cds /cds=(40,423) /gb=U32576 /gi=975892 /ug=Hs.110675 /len=613	34454_r_at
	AL096717				Cluster Incl. AL096717:Homo sapiens mRNA; cDNA DKFZp564P0662 (from clone DKFZp564P0662) /cds=UNKNOWN /gb=AL096717 /gi=5419852 /ug=Hs.24178 /len=2228	41328_s_at
C22orf4 (chromosome 22 open reading frame 4)	AL096779	Hs.20017		22q13.3	Cluster Incl. AL096779:Novel human gene mapping to chomosome 2213.3 similar to yeast ORF YOR070C, putative GTPase Activator (start missing) /cds=(51,917) /gb=AL096779 /gi=5420221 /ug=Hs.20017	33778_at

					/len=1416	
KIAA0257(RW1 protein)	D87446	Hs.75912		2	Cluster Incl. D87446:Human mRNA for KIAA0257 gene, partial cds /cds=(0,5418) /gb=D87446 /gi=1665780 /ug=Hs.75912 /len=6178	36971_at
KIAA0793(KIAA0793 gene product	AB018336	Hs.301283	NM_014808	2	Cluster Incl. AB018336:Homo sapiens mRNA for KIAA0793 protein, complete cds /cds=(117,3281) /gb=AB018336 /gi=3882306 /ug=Hs.26885 /len=3997	35188_at
IL24 (interleukin 24)	AA214546	Hs.315463	NM_006850	1q32	Cluster Incl. AA214546:zr92c03.s1 Homo sapiens cDNA, 3 end /clone=IMAGE-683140 /clone_end=3 /gb=AA214546 /gi=1813171 /ug=Hs.66576 /len=516	41847_at
CSTF3 (cleavage stimulation factor, 3' pre-RNA, subunit 3, 77kD)	U15782	Hs.180034	NM_001326	11	Cluster Incl. U15782:Human cleavage stimulation factor 77kDa subunit mRNA, complete cds /cds=(131,2284) /gb=U15782 /gi=632497 /ug=Hs.180034	41183_at

					/len=2766	
	AB002448				Cluster Incl. AB002448:Homo sapiens mRNA from chromosome 5q21-22, clone-357Ex /cds=UNKNOWN /gb=AB002448 /gi=2943811 /ug=Hs.26968 /len=1270	36260_at
BICD1 (Bicaudal D (Drosophila) homolog 1)	U90028	Hs.164975	NM_001714	12p11.2-p11.1	Cluster Incl. U90028:Homo sapiens bicaudal-D (BICD) mRNA, complete cds /cds=(81,3008) /gb=U90028 /gi=2745975 /ug=Hs.164975 /len=3257	40548_at
GTT1(GTT1 protein)	AL041780	Hs.283722	NM_020151	2	Cluster Incl. AL041780:DKFZp434A0418_s1 Homo sapiens cDNA, 3 end /clone=DKFZp434A0418 /clone_end=3 /gb=AL041780 /gi=5421127 /ug=Hs.239060 /len=723	41295_at

	AL03674				Cluster Incl. AL036744:DKFZp564I1663_r1 Homo sapiens cDNA, 5 end /clone=DKFZp564I1663 /clone_end=5 /gb=AL036744 /gi=5927888 /ug=Hs.236327 /len=617	41288_at
SLC29A2 (solute carrier family 29 (nucleoside transporters), member 2)	AF034102	Hs.32951	NM_001532	11q13	Cluster Incl. AF034102:Homo sapiens NBMPR-insensitive nucleoside transporter ei (ENT2) mRNA, complete cds /cds=(237,1607) /gb=AF034102 /gi=2811136 /ug=Hs.32951 /len=2522	39661_s_at
KIAA0769(KIAA0769 gene produc	AB018312	Hs.19056	NM_014824	11	Cluster Incl. AB018312:Homo sapiens mRNA for KIAA0769 protein, complete cds /cds=(239,2293) /gb=AB018312 /gi=3882258 /ug=Hs.19056 /len=4326	32224_at
KYNU (kynureninase (L-kynurenine hydrolase)	U57721	Hs.169139	NM_003937	2p14-q21.3	Cluster Incl. U57721:Human L-kynurenine hydrolase mRNA, complete cds /cds=(106,1503) /gb=U57721 /gi=1323714	40672_at